UNIVERSITÉ DE MONTRÉAL

**CENTRE DE RECHERCHE SUR LES TRANSPORTS**

CASE POSTALE 6128, SUCCURSALE A, MONTRÉAL, QUÉBEC, H3C 3J7

# PLANIFICATION D'UNE FLOTTE D'AVIONS

par
François Soumis

Université de Montréal

Centre de recherche sur les transports - Publication # 133
Département d'informatique et de recherche opérationnelle
- Publication # 318

Décembre 1978

Dépôt légal - 2e trimestre 1979

# TABLE DES MATIERES


| | Page |
|---|---|
| TABLE DES MATIERES | i |
| TABLE DES ILLUSTRATIONS | v |
| SOMMAIRE | viii |
| CHAPITRE 1: INTRODUCTION | 1 |
| 1.1 Situation du problème dans le processus de planification | 2 |
| 1.1.1 Vocabulaire | 2 |
| 1.1.2 Classification offre/demande | 3 |
| 1.1.3 Classification selon le niveau de planification | 6 |
| 1.1.4 Diagramme général du problème | 10 |
| 1.2 Revue de la littérature | 13 |
| 1.2.1 Modèles microscopiques | 14 |
| 1.2.2 Modèles intermédiaires | 15 |
| 1.2.3 Modèles détaillés | 16 |
| 1.3 Orientation choisie pour formuler le problème | 17 |
| 1.3.1 Approche globale | 18 |
| 1.3.2 Optimisation égoïste versus optimisation coopérative | 19 |
| 1.3.3 Formulation comme un problème d'optimisation | 21 |
| CHAPITRE 2: MODELES DE REPARTITION DES PASSAGERS | 26 |
| 2.1 Réseau utilisé par les passagers | 27 |
| 2.1.1 Topologie du réseau | 27 |
| 2.1.2 Notations | 32 |
| 2.2 Limite de capacité des avions | 33 |
| 2.2.1 Notion de passagers excédentaires | 34 |
| 2.2.2 Calcul du nombre de passagers excédentaires par arc | 36 |
| 2.2.3 Calcul du nombre de passagers excédentaires pour une route | 42 |

| | | Page |
|---|---|---|
| 2.3 | Attraction des routes | 44 |
| | 2.3.1 Facteurs et relations utilisés | 44 |
| | 2.3.2 Regroupement des passagers par heure de départ | 46 |
| 2.4 | Modèle P1: Minimisation de la perte de revenus résultant des passagers excédentaires et de l'utilisation de routes non attrayantes | 47 |
| | 2.4.1 Perte de revenus résultant des passagers excédentaires | 48 |
| | 2.4.2 Coût associé à l'utilisation de routes non attrayantes | 49 |
| | 2.4.3 Unicité et stabilité de la solution | 51 |
| 2.5 | Modèle P2: Compromis entre la perte de revenus due aux passagers excédentaires et une répartition en proportion de l'attraction des routes | 55 |
| | 2.5.1 Comportement des passagers en fonction de l'attraction des routes | 55 |
| | 2.5.2 Evaluation des coefficients d'attraction | 56 |
| | 2.5.3 Objectif correspondant à ce comportement des passagers | 57 |
| | 2.5.4 Stabilité de la solution | 60 |
| 2.6 | Modèle P3: Répartition des passagers en considérant l'"optimisation égoïste" | 61 |
| | 2.6.1 Comportement des passagers | 63 |
| | 2.6.2 Equations du modèle | 66 |
| | 2.6.3 Propriétés du modèle | 69 |
| 2.7 | Modèle P4: Répartition en considérant l'"optimisation égoïste" et une demande élastique | 71 |
| | 2.7.1 Mesure de la disponibilité des routes | 72 |
| | 2.7.2 Relation entre le nombre de passagers et la disponibilité des routes | 73 |
| | 2.7.3 Equations du modèle | 75 |
| | 2.7.4 Propriétés du modèle | 76 |

| | | Page |
|---|---|---|
| CHAPITRE 3: ALGORITHME DE RESOLUTION DU PROBLEME DES PASSAGERS | | 78 |
| 3.1 | Décomposition du problème | 80 |
| 3.1.1 | Convergence de la méthode de décomposition pour les modèles P1 et P2 | 81 |
| 3.1.2 | Existence de solutions du problème P3 | 88 |
| 3.1.3 | Existence de solutions du problème P4 | 90 |
| 3.1.4 | Convergence de la méthode de décomposition pour le modèle P3 | 93 |
| 3.1.5 | Convergence de la méthode de décomposition pour le modèle P4 | 101 |
| 3.1.6 | Relaxation de la méthode de décomposition | 102 |
| 3.2 | Algorithmes de résolution des sous-problèmes | 104 |
| 3.2.1 | Méthode de direction réalisable | 107 |
| 3.2.2 | Calcul du pas optimal | 108 |
| 3.2.3 | Simplexe convexe | 110 |
| 3.2.4 | Gradient réduit | 111 |
| 3.2.5 | Méthode du gradient | 111 |
| 3.2.6 | Méthode du second ordre | 117 |
| 3.2.7 | Adaptation des algorithmes pour traiter la demande variable | 132 |
| 3.2.8 | Comparaison des temps de calcul | 135 |
| CHAPITRE 4: EXPERIMENTATION DES MODELES DE REPARTITION DES PASSAGERS | | 137 |
| 4.1 | Plan d'expérience | 137 |
| 4.1.1 | Description du réseau utilisé | 138 |
| 4.1.2 | Calibration du modèle | 141 |
| 4.2 | Validation du modèle avec la demande d'une saison passée | 143 |
| 4.2.1 | Résultats globaux | 143 |
| 4.2.2 | Résultats pour le sous-réseau de l'Ouest Canadien | 148 |
| 4.2.3 | Résultats détaillés pour certaines origines-destinations | 148 |

Page

4.3 Utilisation du modèle avec une estimation de la demande 158

CHAPITRE 5: MODELES DE PLANIFICATION D'UN RESEAU AERIEN 165

5.1 Réseau utilisé par les avions 165

5.2 Structures des modèles proposés 168

5.2.1 Notation 169

5.2.2 Modèles classiques 170

5.2.3 Modèles proposés 173

5.3 Formulation des modèles proposés 174

5.3.1 Modèles G1 et G2: planification d'un réseau aérien en traitant la circulation des passagers suivant l'optimisation coopérative 174

5.3.2 Modèle G3: planification d'un réseau aérien en traitant la circulation des passagers suivant l'optimisation égoïste 176

5.3.3 Modèle G4: planification d'un réseau aérien en considérant une demande variable et la circulation des passagers suivant l'optimisation égoïste 176

5.4 Méthode de résolution 177

5.4.1 Méthodes de Frank-Wolfe 178

5.4.2 Modifications apportées 179

5.4.3 Algorithmes 183

5.4.4 Adaptation de la méthode de résolution pour les modèles G3 et G4

CHAPITRE 6: CONCLUSION 187

ANNEXE 1: Validation des hypothèses utilisées 191

ANNEXE 2: Etude de la convexité de la fonction économique 196

ANNEXE 3: Nombre de passagers excédentaires exprimé en fonction du nombre de passagers trouvant place dans l'avion 207

BIBLIOGRAPHIE 213

# TABLE DES ILLUSTRATIONS


| FIGURES | | PAGES |
|---|---|---|
| 1 | Décomposition du problème suivant les composantes offre et demande | 4 |
| 2 | Problème de planification du service offert par une flotte d'avions | 11 |
| 3 | Réseau associé à un horaire | 29 |
| 4 | Accès des passagers aux routes pour une origine-destination | 31 |
| 5 | Passagers excédentaires en fonction du nombre de passagers assignés pour les différentes approches | 35 |
| 6 | Partage en marché des passagers d'une O-D | 46 |
| 7 | Coût associé à l'inattraction des routes | 50 |
| 8 | Exemple de répartition des passagers avec le modèle P1 | 54 |
| 9 | Exemple de répartition des passagers avec le modèle P2 | 61 |
| 10 | Cheminement des passagers | 64 |
| 11 | Cheminement des passagers suivant l'hypothèse oubliante | 65 |
| 12 | Cheminement des passagers quand une partie des passagers refusés, faute de place, choisit de ne pas voyager | 74 |
| 13 | Graphique de g(u) de la direction du gradient | 115 |
| 14 | Graphique de g(u) de la direction du second ordre | 119 |
| 15 | Valeur de l'erreur en fonction du pas | 131 |
| 16 | Réseau d'Air Canada | 139 |
| 17 | Nombre de passagers estimé par MAPUM en fonction du nombre de passagers observés pour chaque arc | 144 |
| 18 | Distribution de la différence entre l'estimé de MAPUM et le nombre de passagers observés pour chaque arc | 145 |
| 19 | Distribution de la différence entre l'estimé de ASPEM et le nombre de passagers observés pour chaque arc | 145 |

FIGURES


| | | PAGES |
|---|---|---|
| 20 | Réseau de l'Ouest Canadien | 149 |
| 21 | Nombre de passagers estimé par MAPUM en fonction du nombre de passagers observé pour les arcs de l'Ouest Canadien | 150 |
| 22 | Nombre de passagers estimé par ASPEM en fonction du nombre de passagers observé pour les arcs de l'Ouest Canadien | 151 |
| 23 | Résultats détaillés de 76 pour l'O/D YVR - YEG | 152 |
| 24 | Résultats détaillés de 76 pour l'O/D YVR - YYC | 153 |
| 25 | Résultats détaillés de 76 pour l'O/D YEG - YWG | 154 |
| 26 | Résultats détaillés de 76 pour l'O/D YYC - YWG | 155 |
| 27 | Résultats détaillés de 76 pour l'O/D YVR - YWG | 156 |
| 28 | Circulation estimée pour 82-83 pour l'O/D YVR - YEG | 160 |
| 29 | Circulation estimée pour 82-83 pour l'O/D YVR - YYC | 161 |
| 30 | Circulation estimée pour 82-83 pour l'O/D YEG - YWG | 162 |
| 31 | Circulation estimée pour 82-83 pour l'O/D YYC - YWG | 163 |
| 32 | Circulation estimée pour 82-83 pour l'O/D YVR - YWG | 164 |
| 33 | Réseau espace-temps utilisé par un type d'avion | 167 |
| 34 | Réseau agrégé pour un type d'avion | 167 |
| 35 | Modèle classique vs modèle proposé | 172 |
| 36 | Linéarisation de la fontion objective avec les différences finies | 181 |
| 37 | Nombre de passagers par jour par arc comparé à une loi normale | 192 |
| 38 | Réseau pour lequel la fonction économique n'est pas convexe | 203 |
| 39 | Courbe de niveau d'un exemple de fonction économique non-convexe | 205 |

FIGURES


| | | Page |
|---|---|---|
| 40 | Approximation du nombre de passagers refusé en fonction du nombre qui tentent d'utiliser l'avion | 209 |
| 41 | Approximation du nombre de passagers refusés en fonction du nombre qui utilise l'avion | 211 |


TABLEAUX


| | | |
|---|---|---|
| 1 | Approximation pour le calcul du nombre de passagers excédentaires | 41 |
| 2 | Valeur de $E'_a(x_a) - E_a(x_a)/x_a$ | 99 |
| 3 | Caractéristiques des directions de descente des algorithmes étudiés | 108 |
| 4 | Comparaison des temps de calcul des algorithmes étudiés | 136 |
| 5 | Effets du paramètre K | 142 |
| 6 | Répartition de l'erreur par sous-ensemble d'arcs | 147 |
| 7 | Ecart entre les distributions d'heures de départ désirées et observées | 157 |
| 8 | Différence entre les nombres de passagers excédentaires calculés en utilisant la loi normale et en utilisant la distribution empirique | 194 |
| 9 | Points illustrant la non-convexité de $E_a(h)$ | 198 |
| 10 | Points illustrant la non-convexité de $Z_1$ | 204 |

## SOMMAIRE

La planification des opérations d'une compagnie aérienne consiste à déterminer les types d'avions, leurs nombres et l'horaire à utiliser pour desservir les différents marchés. C'est un grand problème combinatoire car, des décisions touchant des lieux et des temps différents sont interreliées. Chaque vol dessert simultanément plusieurs escales ainsi que d'autres villes via les correspondances. De plus, les vols commençant dans une ville et ceux s'y terminant doivent être synchronisés pour bien utiliser les avions et permettre les correspondances utiles.

Ce problème demande de considérer la circulation des avions à laquelle sont associés les variables de décision et les coûts. Il demande de plus, de considérer la circulation des passagers pour évaluer les revenus et la qualité du service offert.

Le problème est généralement traité avec des méthodes "ad hoc" basées sur l'expérience des planificateurs. Certains systèmes informatiques sont utilisés pour regrouper les données pertinentes mais, ils ne prennent ni ne suggèrent la décision. Plusieurs auteurs ont vu l'intérêt d'utiliser la programmation mathématique pour traiter ce problème de façon optimale, en considérant les interactions de façon plus globale. Malgré les efforts déployés, les modèles proposés sont peu employés. Pour obtenir des résultats dans de grands réseaux certains font des compromis sur la représentativité du modèle. D'autres, ayant présenté des modèles plus complets, ont pu traiter que de petits réseaux. Tous ces modèles ont en commun une structure de problème les rendant coûteux à traiter. Avec cette structure, il ne semble pas possible de traiter les grands réseaux de façon détaillée.

Ce travail propose un nouveau type de modèles qui ouvre la voie au traitement détaillé des grands réseaux. Cette technologie

s'inspire de celle qui a permis de grands progrès en planification de la circulation urbaine depuis quelques années. Il s'agit de modèles non-linéaires qui représentent mieux la réalité et qui, de plus, diminuent grandement le temps de résolution.

Les développements les plus importants portent sur le traitement de la circulation des passagers. Le traitement de l'aspect stochastique de la demande en introduisant l'espérance du nombre de passagers refusés, faute de place, à remplacé avantageusement l'utilisation de "Load factor" pour modéliser la limite de capacité des avions. L'utilisation de ce concept a permis de formuler des modèles présentant une structure de problèmes faciles à résoudre.

Quatre modèles de répartition de passagers sont proposés. D'abord, un modèle simple utilise la notion de passagers excédentaires pour modéliser la limite de capacité. Cette notion semble importante car, les résultats de ce modèle sont très bons même si les autres facteurs ont été traités de façon simpliste.

Le deuxième modèle introduit en plus, le comportement pluraliste des passagers. Plutôt que de tous tenter d'utiliser la route la plus attrayante, les passagers tentent de se répartir entre les routes en proportion de l'attraction.

Le troisième modèle reproduit en plus, le comportement égoïste des passagers (User optimization). Chaque passager choisit son itinéraire en fonction de sa seule satisfaction sans se préoccuper de celle des autres ni de l'utilisation rationnelle du réseau.

Le dernier modèle intègre de plus, l'équilibre offre/demande. La demande dépend du service, qui dépend de la disponibilité des routes qui dépend de la circulation des passagers.

Les méthodes de résolution reçoivent une attention toute particulière dans ce travail. L'existence et l'unicité de solution sont étudiés pour chacun des modèles. La structure des problèmes permet la résolution par décomposition. La convergence de la méthode de décomposition est montrée pour chaque modèle. Un algorithme très efficace est développé pour la résolution des sous-problèmes. La preuve de sa convergence et l'ordre de cette convergence sont données.

Le premier de ces quatre modèles a été expérimenté sur le réseau d'Air Canada. Il s'agit d'un grand réseau desservant 76 villes avec 329 vols par jour. L'essai sur une saison passée a permis de comparer l'utilisation des avions calculée avec l'utilisation observée. La moyenne des valeurs absolues des erreurs destination sur chaque segment de vol est de onze passagers. Cette précision se compare très avantageusement avec d'autres modèles. Un autre point très important à souligner est un temps de calcul de seulement 20 secondes CPU pour traiter ce grand problème. Les trois autres modèles n'ont pas encore été expérimentés faute de données.

Ces développements pour le problème des passagers permettent de formuler des modèles globaux dans lesquels l'interaction entre la circulation des passagers et la circulation des avions intervient seulement au niveau de l'objectif. Cette structure permet la résolution par décomposition en traitant alternativement la circulation des passagers et la circulation des avions. La première composante évalue les vols offerts en déterminant comment les passagers les utilisent. La seconde composante, en utilisant cette évaluation, modifie la circulation des avions pour augmenter les bénéfices.

Cette méthode permet d'obtenir des solutions entières pour la circulation des avions mais n'est pas nécessairement optimale. Elle devrait toutefois permettre d'améliorer une solution réalisable en considérant l'impact sur les coûts et les revenus dans l'ensemble

du réseau. Un programme d'ordinateur réalisant cette méthode a été développé et expérimenté sur des données fictives. Les temps de calcul sont satisfaisants mais la qualité des solutions produites par cette méthode sous-optimale devra être évaluée pour différents types de réseau.

Beaucoup de travail reste à faire avant d'en arriver à un "logiciel" général de planification d'un réseau aérien. Par contre, la technologie introduite permet des gains importants du point de vue des propriétés théoriques, de la validité des résultats et du temps de calcul. Il devient possible de formuler un modèle global traitant l'essentiel d'un grand problème dans un temps raisonnable. Il reste à développer les méthodes d'estimation des paramètres d'entrées, à calibrer ces modèles et à faire la démonstration de leur utilité dans un éventail de situations pratiques.

# 1. INTRODUCTION

Un transporteur aérien doit planifier la composition de sa flotte et son utilisation de façon à satisfaire et même à promouvoir la demande de voyage. C'est un problème à la fois difficile et vital pour une compagnie. Il est difficile puisqu'il faut déterminer l'horaire des vols et le cheminement de chaque avion; or ceci constitue un problème combinatoire de grande taille à cause du nombre de facteurs, du nombre de contraintes et du nombre de possibilités impliquées. Il est vital car il détermine le produit vendu, son prix de revient, et par conséquent le bénéfice ou le déficit de la compagnie. La planification devient un problème de plus en plus critique avec la crise énergitique et la situation économique actuelle. D'un côté les coûts d'opérations augmentent rapidement, de l'autre la clientèle s'accroît beaucoup plus lentement ou même diminue, ce qui réduit dangereusement la possibilité de bénéfices. Une telle conjecture demande la revision en profondeur de l'utilisation de la flotte. Ainsi, contrairement aux conditions de la décennie passée où il suffisait d'organiser une croissance régulière dans un marché où la rentabilité était bonne, maintenant il faut restructurer régulièrement le service offert dans un marché beaucoup plus difficile.

Dans ces conditions, des outils utilisant les techniques de pointe de la recherche opérationnelle peuvent rendre de grands services aux gestionnaires. Par contre, dans ce domaine la recherche opérationnelle n'a pas toujours répondu aux attentes, et ce n'est pas durant une telle période critique que les dirigeants d'entreprise vont implanter le premier modèle venu. A cause de ces réticences, il faut un modèle appuyant ses recommandations sur une reproduction suffisamment détaillée du fonctionnement de la compagnie pour permettre aux administrateurs de se convaincre de la valeur de ces dernières.

## 1.1 Situation du problème dans le processus de planification

Après avoir précisé le vocabulaire nous classifions les étapes du processus de planification afin de situer le problème traité.

Les classifications utilisées sont celles que l'on retrouve dans la thèse de Chan (74) qui a traité à peu de chose près le même problème.

Une première classification distingue entre les modèles d'offre de service et les modèles de demande. Une deuxième situe par rapport aux niveaux de planification que l'on retrouve dans le processus de décision d'une compagnie.

### 1.1.1 Vocabulaire

Définissons d'abord certains termes qui reviendront tout au long de ce travail.

*liaison:* Description préliminaire d'un service offert. Ses composantes sont l'itinéraire (liste des aéroports desservis) et le nombre de sièges.

*vol:* Description plus complète d'un service. C'est une liaison pour laquelle on précise en plus les jours d'opération, la fréquence, les types d'avions et les heures d'opération.

*plan de fréquence:* C'est un ensemble de vols desservant un ensemble d'aéroports. Les heures d'opération ne sont pas définitives mais peuvent encore être modifiées.

*horaire:* Plan de fréquence avec des heures d'opération des vols qui sont fixées.

*arc:* Segment entre deux escales d'un vol.

### 1.1.2 Classification offre/demande

Les concepts économiques d'offre et de demande sont à la mode dans l'analyse des transports. Les problèmes d'horaires et de cheminement des avions peuvent être décomposés suivant ces concepts. La figure 1 qui a été proposée entre autre, par Simpson (69) illustre cette décomposition.

Le problème de la demande consiste à évaluer la clientèle entre chaque paire origine-destination. On utilise généralement les méthodes économétriques et des techniques telles que la régression multiple. Les modèles les plus évolués considèrent la relation entre la demande et le niveau de service afin de déterminer la fraction des clients potentiels qui vont voyager avec la compagnie en fonction de l'amélioration ou la dégradation du service.

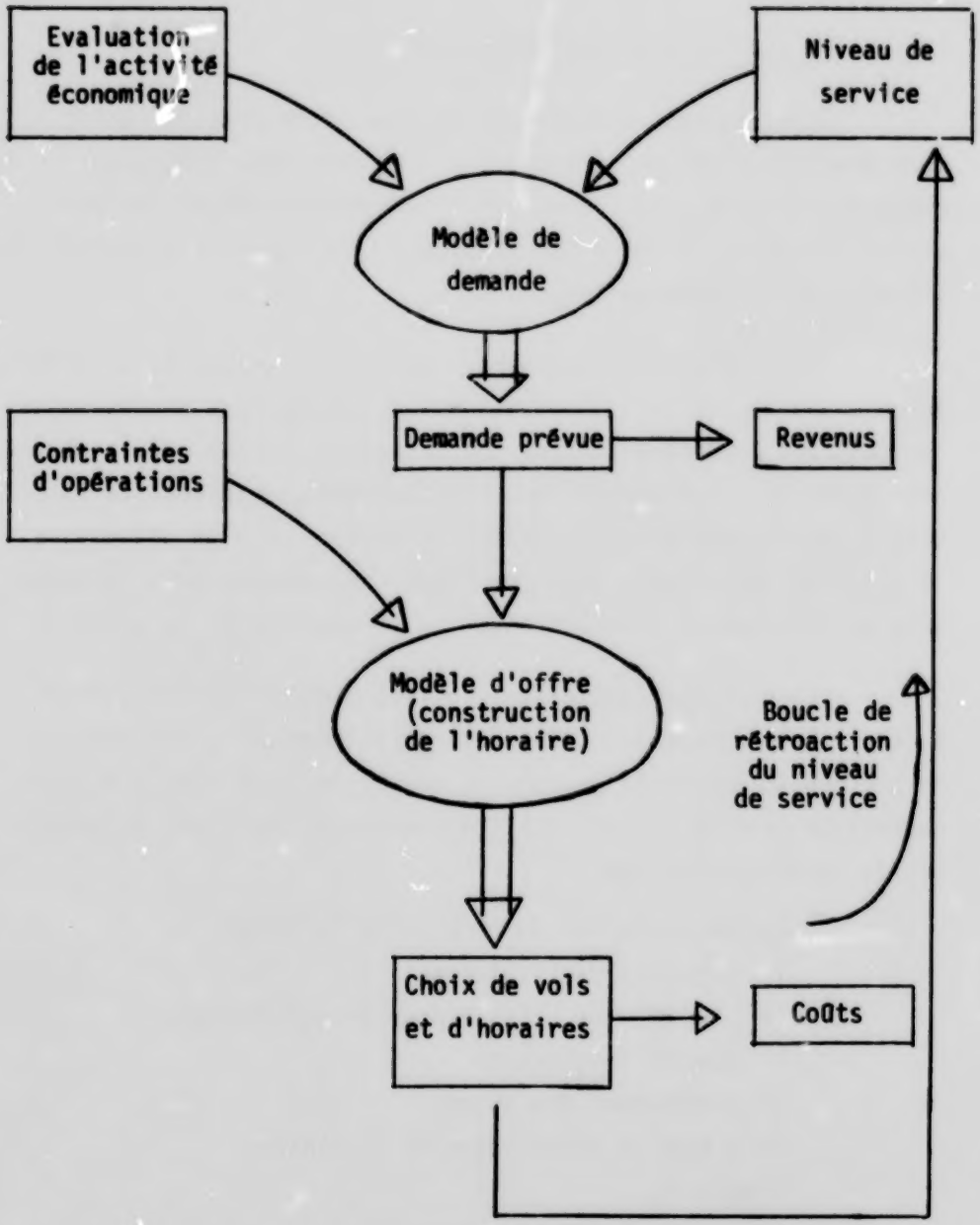
Le problème de l'offre consiste à déterminer le service à offrir pour répondre a cette demande de même que l'utilisation du personnel, des installations et des avions pour offrir économiquement ce service. Voici a titre d'exemples certaines composantes de ce vaste problème:

- la configuration du réseau des liaisons;
- la fréquence des liaisons;
- le type d'avion utilisé pour chaque liaison;
- le plan de fréquence;
- le cheminement des avions;
- la taille et composition de la flotte;
- l'horaire;
- les horaires d'équipages;
- l'assignation des passagers aux vols.

FIGURE 1

DECOMPOSITION DU PROBLEME SUIVANT CES COMPOSANTES

OFFRE ET DEMANDE



Ce problème demande de considérer simultanément deux réseaux dans l'espace et le temps. Un réseau de vols dans lequel cheminent les avions et un réseau d'arcs sur lesquels voyagent les passagers.

Généralement ce problème est traité avec une demande fixe. Ceci conduit a des solutions non-équilibrées du point de vue offre/demande. Des versions plus sophistiquées de modèles d'offre prévoient une boucle de rétroaction (feed-back) entre l'offre et la demande. Ainsi le niveau de service déterminé par le modèle d'offre sert a réajuster la demande dans un modèle considérant ce facteur. Ce processus peut permettre d'approcher une solution d'équilibre du point de vue offre/demande.

Dans ce travail nous proposons d'abord des modèles d'offre utilisant une demande fixe et évaluée au préalable. Les modèles d'assignation de passagers présentés aux sections 2.4, 2.5, 2.6 et les systèmes d'horaire et de cheminement des avions présentés aux sections 5.3.1, 5.3.2 sont de ce type. Même s'ils sont incomplets, ces modèles sont intéressants car ils demandent moins de données et s'intègrent facilement dans le processus actuel de planification qui étudie séparément la demande et l'offre. Un modèle de demande qui tient compte de l'influence du niveau de service peut être associé à ces modèles d'offre pour former une boucle de rétroaction permettant d'approcher la solution d'équilibre du point de vue offre/demande. Le développement de nouveaux modèles informatisés, précis et rapides facilite l'exploitation de cette boucle de rétroaction.

Nous introduisons ensuite un nouveau type de modèle qui traite simultanément la demande et l'offre. Il ne s'agit pas d'une boucle de rétroaction mais de modèles d'équilibre. L'estimation de la demande potentielle doit être faite au préalable en considérant les facteurs économiques. Par la suite, on traite simultané-

ment dans un même problème l'offre et la relation entre le niveau de service et la partie de la demande voyageant avec la compagnie. Le modèle d'assignation de passagers, présenté à la section 2.7 et le système d'horaire et de cheminement des avions présenté à la section 5.3.3 sont des modèles d'équilibre.

### 1.1.3 Classification selon le niveau de planification

Une caractéristique propre de presque tous les problèmes de planification est que certaines décisions doivent être prises assez longtemps à l'avance, tandis que les projections nécessaires sont de moins en moins fiables à mesure que l'horizon s'éloigne. Face à ce dilemme, la solution utilisée est de planifier par niveau. On prend d'abord les décisions d'orientation à long terme basées sur la tendance générale des prévisions. Par la suite, à mesure que les prévisions détaillées deviennent plus fiables, la planification peut descendre au niveau des détails.

Malgré cette stratification, le problème reste difficile car ces niveaux sont très dépendants. L'intérêt d'une décision à un niveau supérieur dépend grandement de son impact sur les niveaux inférieurs. Par exemple, pour choisir les types d'avions à acheter, il faut considérer le détail de leur utilisation au cours des prochaines années. En effet, les types d'avions ne peuvent être choisis en considérant seulement le coût de revient par kilomètre-passager. Un type d'avion supérieur de ce point de vue n'est pas nécessairement le plus rentable pour un réseau donné. En effet, les appareils ayant un faible coût d'opération par siège sont de gros appareils qui fonctionneraient à perte dans plusieurs marchés faute de remplissage. De même, chaque liaison ne peut être considérée séparément. Un type d'avion idéal pour une liaison pourra être rejeté si la géographie du réseau ne per-

met pas d'utiliser ces avions de façon rentable toute la journée.

Afin de situer le problème traité, nous décrivons maintenant les différents niveaux de planification avec leurs principales variables de décision, l'information disponible et l'échelle de temps. Etant donné la forte dépendance entre les différents niveaux il ne faut pas y voir une suite de problèmes qui sont résolus séquentiellement. Il s'agit d'un même problème qui est résolu à des moments différents avec des niveaux d'agrégations différents. Pour la planification à moyen terme l'effort est concentré sur les variables à fixer longtemps à l'avance, en approximant de la meilleure façon possible l'effet des niveaux inférieurs. Pour le plus court terme les variables supérieures sont déjà fixées et l'effort porte sur les variables de plus bas niveau.

La planification à moyen terme porte sur les politiques de la compagnie. Il s'agit de déterminer l'évolution de la flotte et du service à offrir pour chaque marché.

La composition de la flotte est exprimée en nombre d'avions de chaque type. L'évolution doit se faire de façon continue car on ne peut pas acheter et revendre à chaque année des avions pour constituer la flotte idéale. Seuls, certains types d'avions sont encore disponibles sur le marché tandis que la revente est peu profitable. De plus, il faut limiter le nombre de types d'avions pour ne pas augmenter indûment les coûts d'entretien. Les décisions sont donc: quels nouveaux types d'avions introduire, à quel rythme, quand et à quel rythme retirer les avions plus anciens.

Le service à offrir est déterminé en terme de liaison. Ici encore l'évolution doit se faire de façon continue pour des raisons de marketing. Le service peut être amélioré en augmentant le nombre de sièges, en diminuant le nombre d'escales et en augmentant la fréquence, en créant de nouvelles liaisons pour remplacer des vols avec correspondances, etc.

Ces décisions doivent être prises jusqu'à 3-4 ans à l'avance à cause des délais de livraison des avions, des délais de négociation avec les gouvernements et les autres compagnies pour le partage des marchés et finalement pour permettre de planifier l'organisation au sol dans chaque ville. L'information utilisée à ce niveau comprend:

- les projections de la demande entre chaque paire de villes;
- les revenus par passager pour chaque paire de villes;
- les coûts d'opération de chaque type d'avion;
- des propositions de modifications de liaison préparées par le marketing;
- le coût du financement et sa disponibilité.

Les principales relations que l'on retrouve à ce niveau peuvent être illustrées en discutant des effets possibles de l'augmentation de la demande sur la composition de la flotte et le plan de fréquence. Ainsi une augmentation de la clientèle sur le réseau long courrier nécessite l'addition de transporteurs géants. Par contre, une même augmentation sur les liaisons de courte distance peut se traduire par une augmentation de la fréquence et justifier l'acquisition de petits avions. D'autre part, si la fréquence est suffisante, on peut remplacer ces petits avions par des plus gros. On peut également diminuer le nombre d'escales tout en conservant des fréquences suffisantes pour les villes intermédiaires et ceci nécessite des avions à plus long rayon d'action.

La planification à court terme se fait au niveau de l'organisation annuelle. Il s'agit de déterminer les différents plans de fréquence pour s'adapter aux variations saisonnières de la demande. Etant donné la flotte disponible, il s'agit de déterminer les fréquences, l'horaire et les types d'avions à utiliser pour offrir les liaisons désirées. Les différents plans de fréquence des saisons consécutives ont avantage à avoir le plus grand nombre d'éléments en commun pour faciliter la transition.

L'information nécessaire est de même nature qu'au niveau précédent. Par contre, afin de prendre des décisions fiables au niveau de chaque vol, on a besoin d'estimations beaucoup plus précises de la demande.

Pour évaluer les bénéfices attendus d'un plan de fréquence, il faut d'une part évaluer la clientèle attirée par chacun des vols afin de déterminer les revenus. D'autre part, il faut déterminer le cheminement des avions pour évaluer les coûts d'opération. De plus, il faut s'assurer que la solution est réalisable. C'est-à-dire que les passagers doivent trouver une place sur une route acceptable à une heure ne s'écartant pas trop de l'heure désirée, et que le cheminement des avions est possible avec la flotte disponible.

La planification à très court terme se situe au niveau du contrôle des opérations. Il s'agit de déterminer l'horaire exact des vols et l'appareil qui effectue chacun des vols. Cette étape part du plan de fréquence pour fabriquer un horaire qui tient compte de tout le détail des opérations. L'horaire est déterminé avant chaque saison et ne sera modifié qu'en situation d'urgence. Par contre, le choix de l'appareil et même le type d'appareil peuvent être modifiés pour des raisons d'entretien et pour faire face aux pannes et même pour s'adapter à une demande différente de celle prévue.

Les données du problème comprennent:

- le plan de fréquence;
- les contraintes d'aéroports; le nombre de portes par type d'avion, les délais entre les atterrissages;
- les contraintes d'entretien; la fréquence, le lieu et la durée des vérifications par type d'avion;
- le temps au sol; le temps de préparation entre les vols, le temps d'escale par type d'avion et par aéroport.

Ces contraintes ont déjà été partiellement considérées au niveau précédent. Il s'agit ici de faire des modifications mineures à l'horaire pour satisfaire les contraintes négligées.

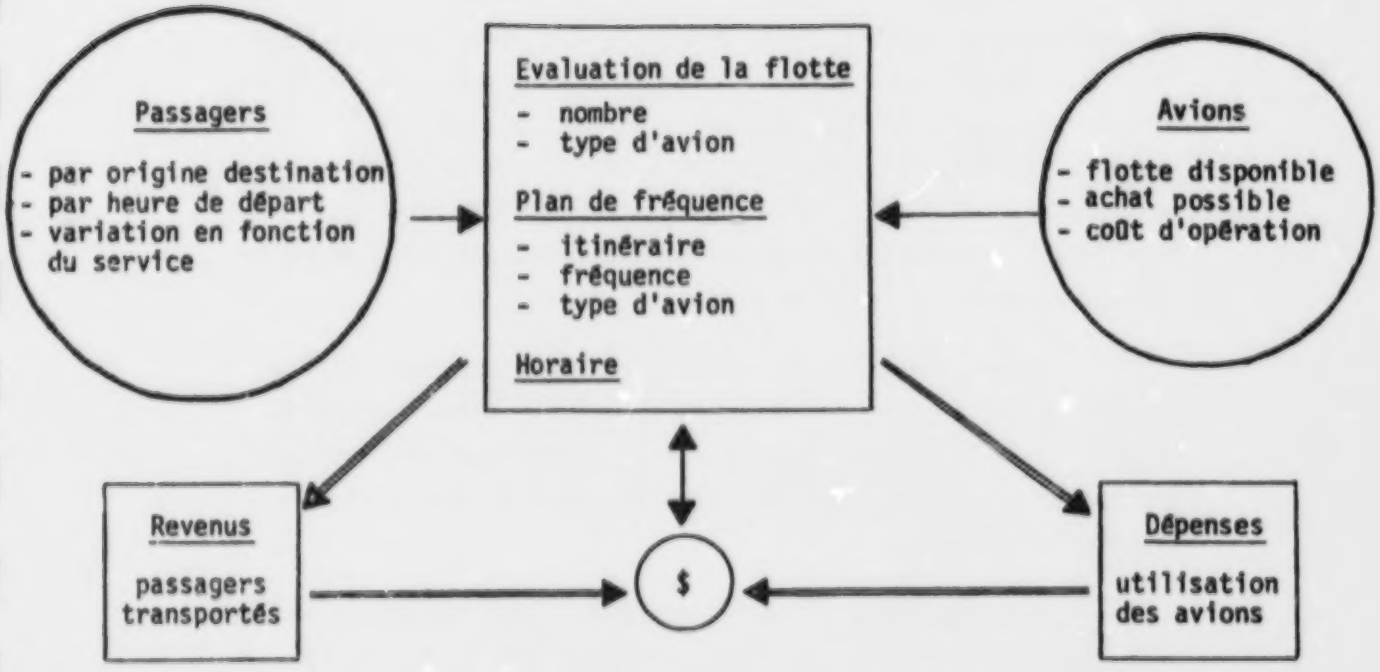
Les outils que nous proposons ne sont pas conçus pour la planification à très court terme. Ce niveau sera seulement approximé lors du traitement des niveaux supérieurs. Les modèles développés sont adaptés pour la planification à court terme. Ils se classent parmi les modèles les plus détaillés et les plus complets jamais proposés pour traiter ce niveau de planification. D'autre part, il peuvent aussi être utilisés pour étudier l'effet des politiques à moyen terme sur le fonctionnement et la rentabilité des opérations. Cette dernière utilisation des modèles est facilitée par le fait qu'ils ont des temps d'exécution très courts qui permettent de réaliser plusieurs simulations à peu de frais dans un court laps de temps.

### 1.1.4 Diagramme général du problème

Nous allons maintenant clarifier les idées en représentant un diagramme du problème traité. Il s'agit du problème de l'offre et de la partie du problème de la demande qui est influencée par l'offre. Du point de vue des niveaux de planification, nous considérons les deux premiers niveaux qui, comme nous l'avons vu, demandent de résoudre le même problème en adaptant le niveau de détail.

FIGURE 2

PROBLEME DE PLANIFICATION DU
SERVICE OFFERT PAR UNE FLOTTE D'AVIONS



Les données du problème sont encerclées. Du côté des passagers, nous considérons la demande par paire de villes en fonction de l'heure de départ désirée, la flexibilité relativement à l'heure de départ et finalement la variation de la demande en fonction du service offert. Du côté des avions, nous considérons la flotte disponible, les achats et retraits possibles et les coûts d'opération. De plus, nous considérons la réglementation gouvernementale, l'autonomie des avions et la limite des installations de certains aéroports.

Le rectangle central contient les variables des décisions. Elles apparaissent dans l'ordre suivant les niveaux de planification. La planification à moyen terme considère les variations possibles de la flotte et des itinéraires. A ce niveau, on considère de façon implicite les variables des niveaux inférieurs pour en évaluer l'impact au niveau supérieur. La planification à court terme considère la flotte comme fixe et étudie les variations du plan de fréquence. On considère l'opportunité d'ajouter ou d'enlever des vols, de changer de type d'avion, de modifier l'horaire, etc.

Pour déterminer la rentabilité d'une solution, il faut évaluer les revenus et les dépenses impliqués. Pour évaluer le revenu, il faut estimer le nombre et le cheminement des passagers transportés. Pour évaluer les dépenses, il faut déterminer l'utilisation des avions.

La majorité des modèles proposés jusqu'ici présentent des approximations au niveau des sous-problèmes dans le but de réduire la complexité du problème global. Ainsi au lieu de considérer la circulation des avions dans le réseau pour évaluer les coûts, on les évalue en limitant seulement le nombre moyen d'heures d'utilisation journalière de chaque type d'avion et on considère qu'il sera possible de faire un horaire par la suite.

Une approximation draconienne du côté des revenus consiste à considérer une demande fixe et par conséquent des revenus fixes. Une autre approximation plus courante consiste à supposer que si le nombre moyen de passagers d'un avion est inférieur à un certain coefficient de remplissage, tous ces passagers peuvent être transportés. Nous allons traiter ces deux sous-problèmes dans le détail sans faire de telles approximations. Les revenus et dépenses seront évalués en simulant à l'avance le fonctionnement de l'horaire du point de vue circulation des passagers et des avions.

## 1.2 Revue de la littérature

Les travaux sur ce sujet sont trop nombreux pour qu'il soit possible de tous les présenter ici. Nous allons insister surtout sur ceux qui proposent des approches compatibles avec le diagramme général présenté à la figure 2. Ces travaux ont aussi en commun avec notre approche d'utilisation de la programmation mathématique.

Nous sommes conscients de négliger toute une autre école qui produit des systèmes moins attrayants du point de vue théorique mais largement utilisés. Il s'agit des systèmes qui améliorent une bonne solution trouvée manuellement en utilisant des règles heuristiques. Il est difficile de faire le point sur ce type de systèmes, car la plupart sont propriétés privées et leurs principes n'ont pas été publiés. De plus, les évaluations et les comparaisons sont délicates car leur fonctionnement est souvent très dépendant des conditions de l'application. Leur utilisation tient plus de l'art que de la science. Les techniques de Rubin (71) et Pollack (73) sont des exemples de ce type de méthodologie.

Le diagramme général présenté à la figure 2 a été traité à différents niveaux de détail. Plusieurs approximations ont été

proposées au niveau des sous-problèmes de circulation des passagers et des avions de façon à réduire la taille du problème. Voyons quelques étapes importantes dans le processus de raffinement de cette famille de modèles.

### 1.2.1 Modèles microscopiques

Les premières approches suivant le diagramme général ont été proposées par Elce (70) et Agar, Sudarovick et Hemmer, (73). **Les variables de décisions sont la flotte et le plan de fréquence.**

Les approximations suivantes sont communes à ces travaux.

**Sous-problèmes des passagers**

A1: La demande est indépendante du service

A2: Les nombres de passagers par avion découlent directement de la demande

A3: Le problème stochastique de la limite de capacité d'un avion est remplacé par une contrainte limitant le remplissage moyen de l'avion par un coefficient maximum (maximum load factor)

Sous-problèmes des avions

A4: Les contraintes de circulation des avions dans l'espace et le temps sont remplacées par une limite sur le nombre journalier moyen d'heure d'utilisation de chaque type d'avion.

Cette approximation consiste à supposer que si les vols peuvent être couverts en utilisant les avions qu'un nombre limité d'heures par jour il est possible à posteriori de faire un horaire où les avions assurent tous ces vols.

Le modèle de Elce utilise comme approximation A2, une procédure heuristique pour déterminer les itinéraires des passagers. Agar et autres supposent que ces itinéraires sont prédéterminés et connus. La résolution utilise la programmation linéaire et des procédures d'arrondissement pour rendre les solutions entières.

A cause de leur portée limitée, ces modèles doivent être complétés par des procédures manuelles et par des modèles complémentaires. L'ensemble forme une suite incluant entre autres un modèle d'évaluation de la demande, un modèle de fabrication d'horaire et un modèle d'évaluation financière. Les intéractions entre les éléments de cette suite sont traitées en considérant des boucles de rétroaction.

Le modèle de Elce développé pour Air Canada n'a pas apporté des résultats escomptés à cause de la difficulté de faire l'horaire après avoir déterminé le plan de fréquence. Le modèle de Agar et autres a donné pour le réseau d'Air France des résultats assez satisfaisants à cause des particularités de ce réseau. En effet, la structure étoilée du réseau (presque tous les vols commencent ou se terminent à Paris) regroupe les avions disponibles où ils doivent être utilisés, ce qui rend beaucoup plus facile la fabrication d'horaire à posteriori.

### 1.2.2 Modèles intermédiaires

Les variables de décision sont encore la flotte et le plan de fréquence. L'approximation A2 disparaît et les passagers ont un choix d'itinéraires suivant l'attraction et la disponibilité des avions. Ceci permet une meilleure évaluation de l'utilisation des avions et des revenus associés.

Les modèles d'assignation de passagers présentés par Charnes and Cooper (61) et Price et autres (74) traitent le sous-problème de

la circulation des passagers en utilisant seulement les approximations A1 et A3. Ces modèles sont relativement coûteux à résoudre malgré la structure de réseau sous-jacente car il s'agit d'un flot de plusieurs commodités avec bornes communes sur les arcs.

Le modèle de planification de flotte développé au Flight Transportation Laboratory (Swan, 72) ne conserve que les approximations A3 et A4. La demande entre deux villes varie avec le service. Elle est fonction de la fréquence journalière totale qui est une somme des fréquences pondérées suivant le type d'avion et le nombre d'arrêts. Une variante de ce modèle (Daigle, 76) permet de plus l'utilisation par les passagers d'itinéraires comportant des correspondances.

Ces modèles introduisent des progrès du côté du sous-problème des passagers mais nécessitent toujours la fabrication a posteriori de l'horaire, ce qui est généralement très difficile. De plus, la solution obtenue est très dépendante des limites artificielles imposées par les approximations A3 et A4.

### 1.2.3 Modèles détaillés

Des modèles détaillés, considérant la flotte et l'horaire comme variables de décision ont été proposés par Agin (74), Chan (74), Richardson (76) et Scarbrough (77). Ces modèles considèrent explicitement la circulation des avions et n'utilisent plus l'approximation A4. Les approximations A1 et A3 y sont par contre utilisées.

Ces modèles donnent lieu à de très grands problèmes linéaires qui sont traités par la décomposition de Dantzig-Wolfe où des variantes heuristiques donnent des solutions entières. La circulation des passagers forme le problème maître et la circulation de chacun des types d'avions forme un sous-problème.

Agin s'est limité aux développements théoriques. Chan a tenté de traiter de grands problèmes en utilisant des méthodes efficaces de traitement des réseaux sous-jacents. Le système est très complexe même s'il n'a considéré qu'un type d'avion et imposé des restrictions à l'utilisation par les passagers des routes avec correspondances et escales. Son travail démontre la difficulté de développer un modèle détaillé pour traiter les grands réseaux en utilisant l'approximation A3 et la programmation linéaire. Les modèles de Richardson et Scarbrough utilisent des techniques de la programmation en nombre entier pour trouver une solution entière optimale. Avec ces méthodes, seuls des petits réseaux peuvent être considérés à cause des limites de temps et de mémoire d'ordinateur.

## 1.3 Orientation choisie pour formuler le problème

Après avoir situé le problème à traiter, nous allons maintenant discuter de l'orientation de nos travaux. La ligne directrice a été le développement d'un modèle global traitant le plus simultanément possible l'ensemble du problème. Un grand pas dans cette direction a été réalisé en introduisant une nouvelle technologie. Cette technologie est inspirée des travaux de Dafermos (71), Murchland (69), Nguyen (74), qui ont permis de grands progrès en planification de la circulation urbaine depuis quelques années. Il s'agit de modèles non-linéaires qui représentent mieux la réalité et qui, de plus, diminuent grandement le temps de solution. Nous parlons de techniques et d'approches avant de parler de la formulation du problème car c'est le cheminement que nous avons suivi dans le développement du modèle.
En considérant les performances des techniques disponibles pour résoudre différentes structures de problèmes, nous avons cherché comment mettre en équations la réalité pour obtenir des structures de problèmes se traitant de façon économique.

### 1.3.1 Approche globale

Contrairement aux approches séquentielles plus anciennes, nous traitons le problème globalement. Les approches séquentielles présentent toujours le risque de donner des solutions irréalisables ou loin d'être optimales à cause des interactions négligées. Ce risque est d'autant plus gênant que l'on peut apprécier l'erreur, faute de borne, sur la valeur de la solution optimale.

Nous traitons donc un vaste problème d'optimisation non-linéaire dans lequel interagissent le cheminement des avions dans l'espace et le temps, la répartition des passagers sur les vols en considérant les heures de départ préférées et la capacité des avions. Certains des modèles intègrent en plus l'équilibre entre la demande et le service offert à chaque marché. Le détail des contraintes d'entretien et d'aéroport n'est pas considéré afin de réduire la complexité et puisqu'il est ordinairement possible de les satisfaire a posteriori. L'affectation des équipages est aussi exclue, car c'est un problème très complexe et qui a peu d'interaction avec la planification des avions. Du point de vue approche simultanée, ce n'est pas très loin de ce que Chan (74) appelait une utopie.

La résolution fait appel à presque toutes les techniques de décomposition existantes. C'est vraiment la clef qui permet de résoudre un si grand problème. Les décompositions permettent d'isoler des sous-problèmes ayant des structures particulières pour lesquelles il existe des algorithmes efficaces qui par surcroît donnent des solutions entières. A ces méthodes de programmation mathématique sont adjointes des méthodes heuristiques. Elles servent à trouver rapidement des bonnes solutions initiales. Elles servent de plus à répartir de façon pragmatique l'effort de calcul entre les différents niveaux de la décomposition. Elles servent enfin à trouver de bonnes

solutions en nombres entiers sans compléter l'énumération implicite et prouver l'optimalité.

### 1.3.2 Optimisation égoïste versus optimisation coopérative

Ces concepts correspondent à deux façons d'estimer l'utilisation par la clientèle des vols offerts. Ce sont des approches correspondant à deux hypothèses différentes de comportement des passagers. Ces notions ont été introduites par Wardrop (52) dans le contexte du transport urbain sous les noms de "user optimization" et "system optimization".

L'optimisation égoïste correspond à l'hypothèse de comportement suivante:

> "Chaque passager utilise pour se déplacer d'une origine à une destination la route disponible qui lui convient le mieux."

C'est une hypothèse qui décrit bien le comportement humain dans notre société. Chaque passager fait un choix égoïste mais qui dépend du choix des autres car il doit subir l'égoïsme des autres. De façon plus concrète, chaque passager choisit l'itinéraire qu'il préfère mais doit se contenter d'un deuxième choix si l'avion qu'il veut utiliser est déjà rempli à capacité. Ceci amène à la situation vécue ou pour gagner un peu en satisfaction un passager peut en faire perdre beaucoup à d'autres.

Le problème de cheminement des passagers suivant l'approche "user optimization" a été, jusqu'ici, traité seulement avec des techniques heuristiques, (Gunn, 64), (Gagnon, 67), (Kingley, 68). Les systèmes développés arrivent assez bien à déterminer le meilleur itinéraire pour chaque passager mais sont faibles pour déterminer qui devra prendre un deuxième choix quand la capacité de l'avion n'est pas suffisante.

L'optimisation coopérative a été introduite parce qu'elle se prête mieux à l'utilisation de la programmation mathématique. Elle correspond à l'hypothèse de comportement suivante:

> "Chaque passager utilise pour son déplacement la route qui contribue le mieux à un objectif global de satisfaction de l'ensemble des passagers."

C'est une hypothèse normative où l'objectif global détermine le cheminement des passagers. L'itinéraire d'un passager dépend des préférences des autres passagers et de l'intérêt de la compagnie. Quand la capacité d'un avion n'est pas suffisante on déplace ceux qui perdent le moins en prenant un deuxième choix. Le choix de l'objectif global de satisfaction doit être fait de façon à s'écarter le moins possible de l'approche descriptive. Cet écart dépend de l'importance accordée aux intérêts de la compagnie, de la façon de définir la préférence de route et de la structure du réseau.

Même si elle ne correspond pas tout à fait au comportement des passagers,cette deuxième approche a ses avantages. En utilisant la programmation mathématique il est possible de résoudre de façon optimale le problème posé et de faire des analyses de sensibilité sur l'effet des paramètres. Cette dernière possibilité est très utile pour évaluer l'effet sur les passagers de modifier l'ensemble des vols disponibles.

Notre attitude face aux avantages et inconvénients de ces deux orientations a été de préférer utiliser des techniques de programmation mathématique. Nous considérons que pour modifier de façon profitable l'utilisation des avions, il est essentiel de pouvoir faire de l'analyse de sensibilité.

Les premiers modèles de répartition des passagers présentés aux sections 3.1 et 3.2 correspondent à l'optimisation coopérative et utilisent la programmation mathématique. De plus, nous résolvons ce dilemme en présentant aux sections 3.3 et 3.4 des modèles correspondant à l'optimisation égoïste et se résolvant par programmation mathématique. C'est un pas important car ces modèles ont le double avantage de correspondre au comportement des passagers et de permettre l'analyse de sensibilité. Ce développement est l'équivalent en transport aérien de ce que Dafermos (69) a fait pour le transport routier.

### 1.3.3 Formulation comme un problème d'optimisation

Comme nous l'avons vu précédemment, nous traitons plusieurs variantes du problème. Ainsi, nous allons d'abord introduire la partie commune et ensuite présenter chaque variante avec ses particularités.

Les variables du problème et leur domaine de variation se caractérisent comme suit:

i) $L$ est l'ensemble de tous les vols possibles. Cet ensemble ne comprend que des vols en accord avec les réglementations gouvernementales. Les types d'avions utilisés par chacun de ces vols ont une autonomie suffisante pour parcourir la distance et sont autorisés à atterrir dans les aéroports desservis. Il exclut les vols qui à cause de leur itinéraire ou de leur horaire ne sont pas intéressants pour les passagers ou pour la compagnie. Cet ensemble permet de tenir compte a priori des contraintes complexes de nature légale, technique et esthétique limitant la formation des vols. Une méthodologie pour construire cet ensemble a été développée par Bourque (76) en parallèle avec cette thèse.

ii) V est l'ensemble des vols considérés. Cet ensemble définit le réseau dans lequel peuvent évoluer les avions. Ce sous-ensemble permet de limiter le problème aux vols qui paraissent intéressants du point de vue marketing. Ceci permet aussi de faire une étude locale du réseau en prenant comme ensemble , les vols existants et des alternatives de vols dans la région à étudier.

iii) $v \in V$ dénote un vol typique

iv) $y_v$ est une variable binaire prenant la valeur 1 si ce vol fait partie de la solution et prenant la valeur 0 s'il n'en fait pas partie.

v) $Y = (y_v | v \in V)$ est une solution pour la circulation des avions.

vi) G(Y) est le réseau dans lequel voyagent les passagers. Ce réseau est défini par les vols faisant partie de la solution et par les correspondances qu'ils rendent possibles.

vii) D est la demande par heure pour chaque origine-destination. Notons que dans certaines variantes du problème, la demande sera une fonction du service.

viii) H est une solution pour le cheminement des passagers. Nous ne verrons pas maintenant en terme de quelles variables s'exprime cette solution. Mais elles spécifient combien et quels passagers utilisent chaque avion.

Maintenant introduisons les contraintes que doivent satisfaire ces variables. Pour chaque type d'avion, il doit y avoir conservation de flot des avions en chaque ville et ce, en tenant compte des temps des arrivées et des départs. Pour le flot d'avions Y dans le réseau V la conservation de flots est dénotée par la contrainte.

$$F(Y, V) = 0 \qquad (1.1)$$

De même les contraintes de conservation du flot de passagers H dans le réseau G(Y) et la contrainte d'égalité entre le flot et la demande aux origines et destinations sont dénotées par

$$F(H, G(Y)) = D \qquad (1.2)$$

Définissons maintenant les termes qui composent l'objectif

ix) C(Y) est le coût total d'opérations des avions pour les vols faisant partie de la solution

x) R(H) est le revenu associé aux passagers transportés. Les passagers qui ne peuvent être transportés, faute de place, n'apportent pas de revenus.

Voyons maintenant chacune des variantes du problème en commençant avec une formulation avec l'optimisation coopérative en considérant la demande fixe. Si on se contentait de minimiser R(H) - C(Y) sujet aux contraintes (1.1) et (1.2) la solution serait très insatisfaisante. Une telle optimisation du bénéfice de la compagnie forcerait certains passagers à voyager avec une qualité de service inacceptable. La solution obtenue serait irréaliste et très déséquilibrée du point de vue offre/demande. Pour remédier à cette situation on introduit dans l'objectif un terme de satisfaction des passagers, S(H,Y) qui est fonction des vols disponibles et de ceux utilisés.

La formulation devient

$$\begin{aligned} \max \quad & R(H) - C(Y) + S(H, Y) \\ \text{s.a.} \ (1) \quad & F(Y, V) = 0 \qquad\qquad (G1) \text{ et } (G2) \\ (2) \quad & F(H, G(Y)) = D \end{aligned}$$

Cette formulation du problème global est identifiée par (G1) et (G2) car, deux façons différentes de définir S(H, Y) ont été développées. Ces deux variantes seront traitées à la section 5.3.1.

Le sous-problème obtenu lorsqu'on considère un horaire d'avion fixé et que seul le cheminement des passagers peut varier s'énonce comme suit:

$$\begin{aligned} &\max \quad R(H) + S(H, Y) && \text{(P1) et (P2)} \\ &\text{s.a. (2)} \quad F\big(H, G(Y)\big) = D \end{aligned}$$

Les deux variantes (P1) et (P2) correspondant aux deux façons de définir S(H,Y) sont respectivement développées aux sections 2.4 et 2.5.

Pour traiter l'optimisation égoïste, considérons d'abord la situation où la demande est fixe. Ce problème ne se formule pas comme la maximisation d'un objectif global, mais plutôt comme la résolution d'un système d'équations de comportement.

$$E\big(H, S(H, Y), G(Y)\big) = 0 \qquad (1.3)$$

Cette équation relie le cheminement des passagers $H$, l'attrait des vols $S(H, Y)$ et la limite de capacité des avions du réseau $G(Y)$. A ces équations de comportement, on ajoute les contraintes de conservation du flot des passagers pour obtenir le système

$$\begin{aligned} &\text{(2)} \quad F\big(H, G(Y)\big) = D && \text{(P3)} \\ &\text{(3)} \quad E\big(H, S(H,Y), G(Y)\big) = 0 \end{aligned}$$

Ce sous-problème est traité à la section 2.6. Analysons maintenant le problème global correspondant. Etant donné que la résolution du problème de la répartition des passagers est basée sur le comportement de ces derniers plutôt que sur un objectif reflétant le profit de la compagnie, on ne risque pas de sacrifier la qualité du service en tentant d'augmenter le bénéfice. Ainsi, il n'est pas nécessaire d'inclure le terme de satisfaction dans l'objectif qui se réduit à Max $R(H) - C(Y)$ et le problème global correspondant devient

$$
\begin{aligned}
&\text{Max} && R(H) - C(Y) \\
&\text{s.a.} && (1) \quad F(Y, V) = 0 \\
& && (2) \quad F\big(H, G(Y)\big) = D \\
& && (3) \quad E\big(H, S(H, Y), G(Y)\big) = 0
\end{aligned}
\qquad (G3)
$$

Cette variante du problème global est traitée à la section 5.3.2

Comme dernière variante, nous considérons l'approche "user optimization" avec la demande qui varie en fonction du service. Pour définir le sous-problème des passagers il faut introduire le niveau de service N(H, Y) et D(N(H, Y)) une demande fonction du niveau de service

Le sous-problème devient:

$$
\begin{aligned}
&(4) \quad D = D\big(N(H, Y)\big) \\
&(2) \quad F\big(H, G(Y)\big) = D \\
&(3) \quad E\big(H, S(H, Y), G(Y)\big) = 0
\end{aligned}
\qquad (P4)
$$

Ce problème est traité à la section 2.7. Le problème global correspondant s'écrit:

$$
\begin{aligned}
&\text{Max} && R(H) - C(Y) \\
&\text{s.a.} && (1) \quad F(Y, V) = 0 \\
& && (4) \quad D = D\big(N(H, Y)\big) \\
& && (2) \quad F\big(H, G(Y)\big) = D \\
& && (3) \quad E\big(H, S(H, V), G(Y)\big) = 0
\end{aligned}
\qquad (G4)
$$

Cette variante du problème global est traitée à la section 5.3.3. Nous allons maintenant voir au cour des chapitres suivants le détail de ces modèles et les techniques de solutions développées.

## CHAPITRE 2 MODELES DE REPARTITION DE PASSAGERS

Le problème se résume comme suit: étant donné un ensemble de vols, déterminer comment les passagers se déplacent des origines aux destinations, ce, en tenant compte de l'attraction des routes et des heures de vols ainsi que de la limite de capacité des avions. Comme nous l'avons vu ce sous-problème occupe une place importante dans la formulation du problème global. C'est la base de l'évaluation d'un horaire; les améliorations proposées sont d'autant plus valables qu'elles reposent sur une évaluation réaliste de cet horaire. De plus, le problème de répartition des passagers est important en soi.

Il est utile pour l'organisation du service au sol de connaître l'origine-destination des passagers utilisant chaque avion, les embarquements et les débarquements à chaque aéroport, les points de correspondance des passagers. Ce système de répartition des passagers pourrait faciliter la fabrication manuelle d'horaires en évaluant de façon interactive l'effet, sur les passagers, des modifications considérées.

Dans la plupart des systèmes existants, le facteur d'attraction des routes et des heures de vols est traité de façon satisfaisante. Par contre, la contrainte sur la limite de capacité des avions présente beaucoup plus de difficulté. D'une part les méthodes heuristiques employées n'arrivent pas toujours à respecter cette limite de façon acceptable. D'autre part, lorsqu'on considère la limite de capacité des avions sous forme de contraintes explicites et qu'on veut traiter le problème avec des techniques de programmation mathématique, on se retrouve avec un problème très difficile à résoudre puisqu'il s'agit de minimiser le coût d'un flot comportant plusieurs commodités étant donné des bornes sur la quantité totale de flot qui doit passer dans chaque arc.

Dans ce chapitre, nous allons développer une nouvelle façon de traiter la limite de capacité des avions en utilisant la programmation mathématique. Cette façon reflète mieux la réalité tout en diminuant grandement l'effort nécessaire pour résoudre le problème. A partir de cette idée, nous développons dans ce chapitre, 4 modèles en considérant différentes approches. Deux variantes de l'optimisation coopérative se distinguant par la façon de définir l'attraction ainsi que l'optimisation égoiste avec demande fixe et avec demande variable. Commençons par définir le réseau utilisé et les concepts communs aux 4 modèles.

## 2.1 Réseau utilisé par les passagers

A partir de l'ensemble des vols, il faut définir le réseau sur lequel voyagent les passagers. Il s'agit d'un réseau dans l'espace pour représenter la géographie des déplacements. De plus, il faut ajouter la dimension du temps pour distinguer les cheminements différents des passagers voulant voyager aux différentes heures et pour pouvoir représenter les correspondances que l'horaire permet de réaliser. Le problème est de définir un réseau espace-temps le plus petit possible permettant de représenter le détail des déplacements des passagers.

### 2.1.1 Topologie du réseau

Définissons d'abord les arcs et les noeuds du réseau. Les arcs correspondent aux segments de vols entre chaque escale. Les noeuds sont les villes auxquelles on a ajouté la dimension temps. Ainsi, à chaque ville sont associés plusieurs noeuds pour la représenter à

différents moments dans le temps. Si on associe le même nombre de noeuds à chaque ville, le nombre de noeuds du réseau devient très grand. Ainsi, si pour 50 villes on associe un noeud à chacune d'elles à tous les 15 minutes pendant 24 heures, on obtient un réseau ayant près de 5 000 noeuds. Pour certaines villes d'où partent des vols à toutes les 3 heures ce niveau de détails est inutile. Par contre, pour certaines grandes villes un noeud à toutes les 15 minutes n'est pas suffisant car plusieurs avions portent ou arrivent dans cet intervalle aux heures de pointes. Ainsi pour avoir une précision suffisante avec un nombre minimal de noeuds, on associe à chaque ville un noeud aux moments de départ et d'arrivée d'un avion. Dans le cas de départs et d'arrivées simultanés, les noeuds sont dédoublés.

Les arcs de réseau sont définis comme suit:

i) Arc défini entre le noeud de départ et le noeud d'arrivée de chaque segment de vol, les arcs correspondant aux segments de vol entre les escales.

ii) Arc d'escale reliant le noeud d'arrivée et le noeud associé au moment où le vol repart d'une ville.

iii) Arc de correspondance reliant le noeud d'arrivée du premier vol au noeud de départ d'un autre vol dans une même ville. Les arcs de correspondances représentent les correspondances possibles qui sont susceptibles d'intéresser certains passagers.

Les arcs d'escales et de correspondances sont traités implicitement et n'apparaissent pas comme tel dans la formulation et la résolution du problème. La figure 3 illustre les noeuds et les arcs définis par cette méthode à partir d'un horaire de vols.

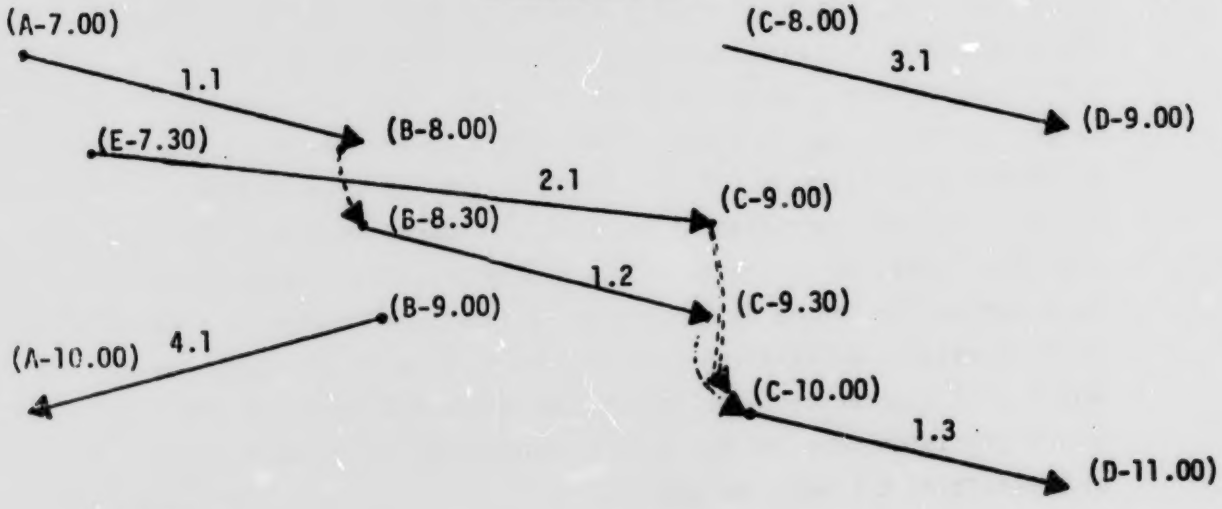
FIGURE 3

RESEAU ASSOCIE A UN HORAIRE

HORAIRE

| Numéro du vol | Ville | Heure d'arrivée | Heure de départ |
|---|---|---|---|
| 1 | A | - | 7.00 |
| 1 | B | 8.00 | 8.30 |
| 1 | C | 9.30 | 10.00 |
| 1 | D | 11.00 | - |
| 2 | E | - | 7.30 |
| 2 | C | 9.00 | - |
| 3 | C | - | 8.00 |
| 3 | D | 9.00 | - |
| 4 | B | - | 9.00 |
| 4 | A | 10.00 | - |

RESEAU ASSOCIE




LEGENDE:

(A-7.00): noeud associé à la ville A à 7.00 heures
: arc d'escale
: arc de correspondance
: arc d'envolée
1.1 : numéro d'arc, (arc 1 du vol 1)

Le problème est traité en terme des routes utilisées par les passagers. D'abord seules les origines-destinations (O - D) ayant une demande journalière moyenne supérieure à un client sont considérées explicitement. Les autres sont décomposées en utilisant un point de correspondance. Considérons pour chacune de ces (O - D) les chaînes utilisables. Une chaîne étant une suite d'arcs qui permet aux passagers de se déplacer de l'origine à la destination en utilisant des escales et des correspondances si nécessaire. Une chaîne est considérée utilisable si certains passagers sont susceptibles de l'utiliser. Les chaînes n'utilisant pas de correspondances sont toutes considérées utilisables. Parmi les chaînes avec correspondances sont excluses celles partant avant une autre chaîne et n'arrivant pas avant. De même sont excluses celles partant en même temps et arrivant plus tard. Ainsi, seule la meilleure correspondance empruntant un arc donné est considérée utilisable par les passagers d'une O-D. Donc, les chaînes desservant une même O-D n'ont pas d'arcs en commun. Le nombre de chaînes utilisables par O-D est en moyenne 3 - 4 pour le réseau d'Air Canada avec un maximum de 30.

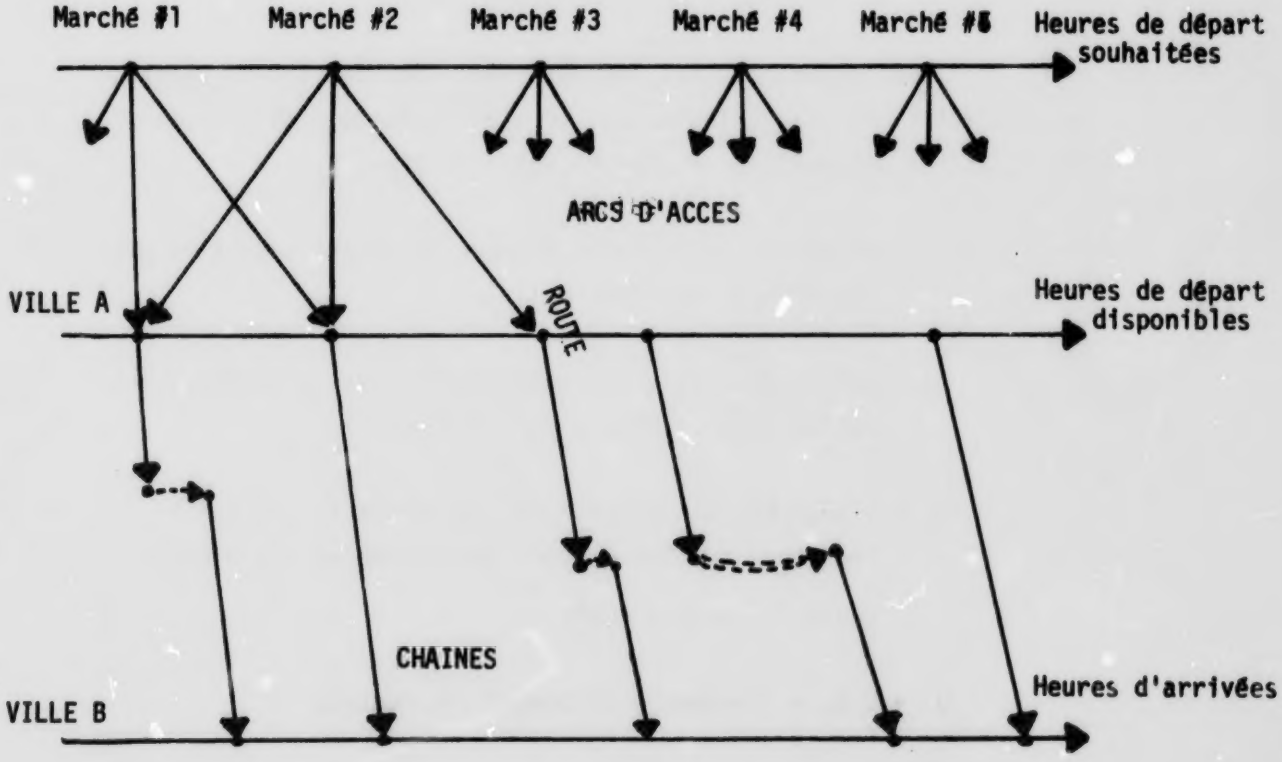
Chaque origine est fractionnée en plusieurs marchés correspondant chacun à un intervalle d'heure de départ désirée. Cette division permet de tenir compte de l'attrait différent des différents vols suivant l'heure de départ désirée. Nous avons choisi de diviser les passagers d'une O-D en autant de marchés qu'il y a de chaînes pour aller de O à D. Chaque marché est associé à une chaîne et défini comme étant l'intervalle d'heures de départ désirées dans lequel cette chaîne est la plus attrayante. Cette façon de déterminer le nombre de marchés a l'avantage de traiter le problème d'une origine-destination avec plus de détails si les heures de départ sont rapprochées. Cette approche donne également une plus grande précision pour les O-D à forte densité où les passagers sont plus sensibles à l'heure de départ.

Définissons maintenant les routes que peuvent utiliser les passagers de chaque marché. Trois routes sont formées en introduisant un arc d'accès du marché vers la chaîne associée, un arc vers la

FIGURE 4

ACCES DES PASSAGERS AUX ROUTES

POUR UNE ORIGINE-DESTINATION



chaîne précédente dans le temps et un vers la suivante. Les deux dernières routes sont des alternatives que peuvent utiliser les passagers si la limite de capacité les empêche d'utiliser la route la plus attrayante. Le choix de définir seulement trois routes par marché a été fait a priori et donne des résultats satisfaisants. De plus, il est possible après la résolution du problème de vérifier si d'autres routes auraient été utiles. Chaque route ainsi définie appartient à un et un seul marché même si les routes de deux marchés peuvent avoir un, plusieurs ou même tous leurs arcs en commun. La figure 4 présente pour une origine-destination le partage en marché et les arcs d'accès définissant les routes.

### 2.1.2 Notations

Nous regroupons ici les notations pour décrire le réseau, les paramètres et les variables qui sont utilisées pour développer les équations des modèles de répartition de passagers.

r = un marché (intervalle d'heure de départ souhaitée pour une origine-destination);

q = une route. Elle est composée d'un arc d'accès, d'une chaîne dans le réseau espace-temps.

$Q_r$ = l'ensemble des routes pour le marché r. Il est à remarquer que les $Q_r$ sont des ensembles disjoints à cause des arcs d'accès;

$Q = \bigcup_r Q_r$ = l'ensemble de toutes les routes;

$r_q$ = le marché desservi par la route q. L'arc d'accès de la route q part de ce marché;

$h_q$ = le nombre de passagers du marché $r_q$ assignés sur la route q;

$h = \{h_q | q \in Q\}$ = une assignation des passagers;

a = un arc de vol. Il correspond au déplacement sans escale d'un avion entre deux villes à une heure donnée;

$E_a$ = l'ensemble des routes utilisant l'arc a;

$F_q$ = l'ensemble des arcs utilisés par la route q;

$f_a$ = le nombre total de passagers assignés à l'arc a

$$f_a = \sum_{q \in E_a} h_q$$

$B^r$ = le coût associé au refus d'un passager du marché r;

$d^r$ = la demande pour le marché r.

## 2.2 La limite de capacité des avions

Un avion est aménagé pour transporter un certain nombre de passagers. Si le vol offert attire plus de passagers que l'avion peut en transporter, ces passagers doivent être détournés sur d'autres vols ou tout simplement refusés. Le modèle est simple quand on considère une seule journée. Par contre, quand on doit traiter simultanément plusieurs journées ayant des nombres variables de passagers, il ne saurait être question d'estimer à l'avance ces nombres pour chaque marché, chaque jour et de faire la planification à ce niveau de détail. Il est d'ailleurs déjà difficile de bien estimer le nombre journalier moyen de passagers désirant voyager. C'est avec ces moyennes que se fait la planification.

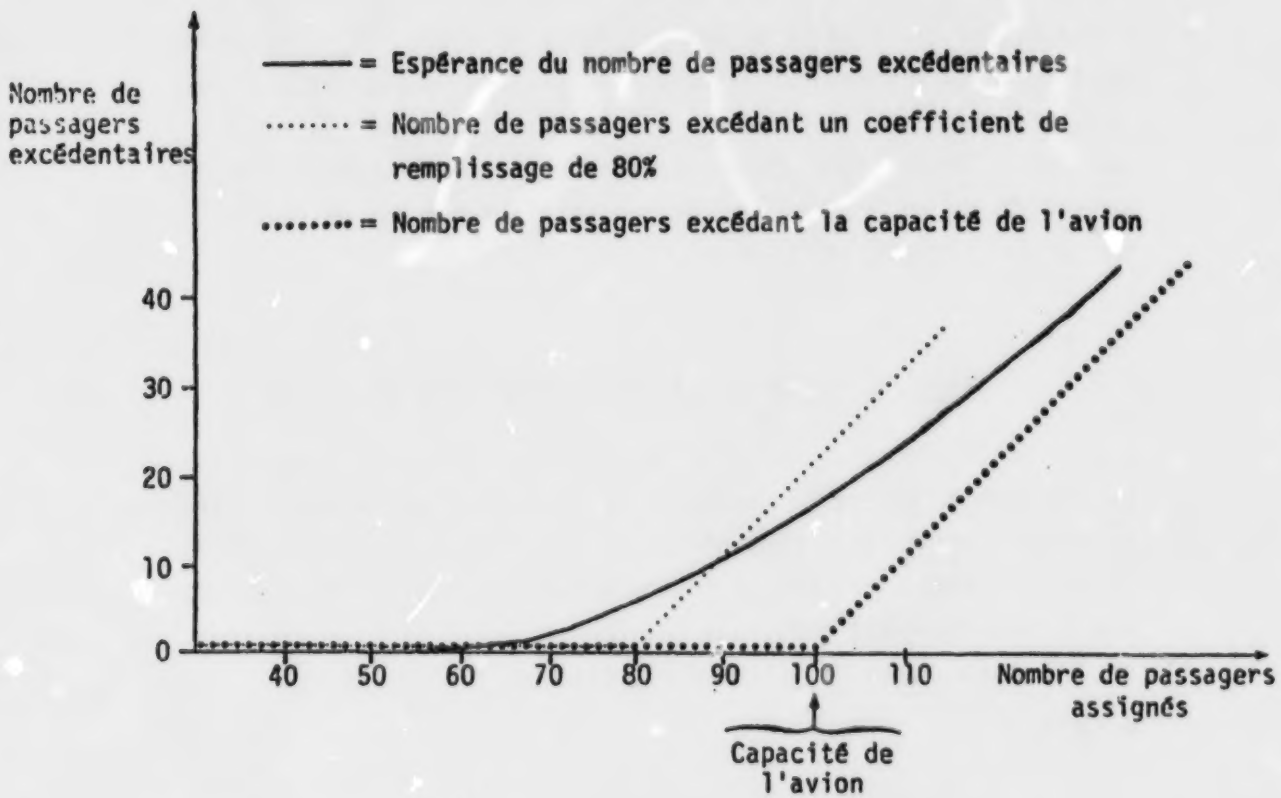
Il ne faut évidemment pas planifier en tentant de remplir les avions en moyenne car il manquerait de place un jour sur deux. Face à ce problème, il est classique de définir un maximum pour le coefficient de remplissage moyen de façon à réserver des sièges pour les fluctuations journalières de la demande. Cette méthode, même si elle est très répandue, présente des problèmes sérieux. Elle est très sensible au choix des coefficients de remplissage. Planifier avec des coefficients plus bas que la valeur idéale se répercute directement sur les coûts d'opération en exigeant plus d'avions, ce qui dilapide une partie importante du bénéfice possible. Par contre, planifier avec des coefficients trop élevés, fait espérer des bénéfices non réalisables.

### 2.2.1 Notion de passagers excédentaires

Comme le même horaire doit se répéter à tous les jours, il n'est pas nécessaire de prévoir le nombre de passagers de chaque jour pour traiter dans le détail le problème de la limite de capacité des avions. Il suffit d'estimer la distribution du nombre journalier de passagers utilisant l'avion. Avec cette distribution du nombre de passagers on peut calculer la probabilité d'occupation des sièges et l'espérance du nombre de passagers excédentaires. Du point de vue de la planification, le nombre de passagers excédentaires est une fonction du nombre moyen de passagers assignés à l'avion. Ce nombre est nul quand le nombre de passagers est faible, devient positif pour les taux de remplissage courants et poursuit sa croissance de façon continue pour atteindre une fraction importante des passagers quand le nombre moyen de passagers assignés dépasse la capacité de l'avion. Ceci diffère de l'approche avec les coefficients maximum de remplissage. En planifiant avec les coefficients de remplissage, on considère que tous les passagers trouvent placé dans l'avion tant que le remplissage moyen n'atteint pas le coefficient maximum, et que tous les passagers additionnels sont refusés faute de place. La figure 5 illustre la différence entre les approches.

FIGURE 5

PASSAGERS EXCEDENTAIRES EN FONCTION DU NOMBRE DE PASSAGERS ASSIGNES POUR LES DIFFERENTES APPROCHES




Suivant l'approche classique les passagers sont assignés sur les routes qui exercent le plus d'attrait en limitant le nombre moyen de passagers par avion de façon à ne pas avoir de passagers

excédentaires. L'assignation des passagers en considérant l'espérance du nombre de passagers excédentaires, donne lieu à un modèle différent car le nombre de passagers excédentaires n'est pas nul pour les niveaux de remplissage courant. L'assignation se fait en considérant que les passagers acceptent de changer de route pour pouvoir voyager si la diminution de satisfaction n'est pas trop importante, sinon ils n'utilisent pas les services offerts par la compagnie pour voyager, ce qui entraîne une perte de revenus pour la compagnie. L'évaluation de la perte de revenus, faute de place, est faite en supposant que le revenu brut associé à un passager est perdu s'il ne peut effectuer son déplacement origine-destination sur la route assignée. On considère qu'un passager ne peut effectuer son déplacement sur la route assignée dès que celui-ci n'a pas de place sur un des tronçons de la route.

### 2.2.2 Calcul du nombre de passagers excédentaires par arc

Il s'agit de calculer le nombre de passagers excédentaires en fonction du nombre de passagers assignés à un avion sur un arc. Ce calcul est fait en considérant la capacité de l'avion utilisé et la distribution du nombre journalier de passagers. Pour permettre un développement théorique ininterrompu, nous utilisons certaines hypothèses dont la validité est discutée à l'appendice 1.

Notation:

$c$ = capacité de l'avion;

$f_a$ = nombre moyen de passagers assignés à l'arc a;

$\sigma_a$ = écart-type du nombre de passagers assignés à l'arc a;

$\lambda$ = taux de variation du nombre journalier de passagers d'un marché;

$E(c, f_a, \sigma_a)$ = espérance du nombre de passagers excédentaires.

$E(c, f_a, \sigma_a)$ se calcule en utilisant la définition de l'espérance:

$$E(c, f_a, \sigma_a) = \int_c^{\infty} (x - c) \; \delta_X(x) \; dx \qquad (2.1)$$

La première hypothèse porte sur la distribution $\delta_X$ que suit le nombre journalier de passagers.

Hypothèse A: Le nombre journalier de passagers suit approximativement une distribution normale.

Cette normalité suggère de réduire le problème du calcul de $E(c, f_a, \sigma_a)$ au cas d'une variable standardisée. Soit:

$$Y = \frac{X - f_a}{\sigma_a} \quad , \quad Y \text{ est } N(0, 1)$$

D'où: $X = \sigma_a Y + f_a$

Exprimons $\delta_X(x)$ en fonction de $\delta_Y(y)$

$$\delta_X(x) = \frac{d}{dx} \; P(X < x)$$

$$= \frac{d}{dx} \; P(\sigma_a Y + f_a < \sigma_a y + f_a)$$

$$= \frac{d}{dx} \; P(Y < y)$$

$$= \frac{d}{dy} \; P(Y < y) \cdot \frac{dy}{dx}$$

$$\delta_X(x) = \delta_Y(y) \cdot \frac{1}{\sigma_a} \qquad (2.2)$$

Effectuons le changement de variable dans la relation (2.1) et utilisons la relation (2.2).

$$\int_c^{\infty} (x - c)\ \delta_X(x)\ dx$$

$$= \int_{\frac{c-f_a}{\sigma_a}}^{\infty} (\sigma_a y + f_a - c)\ \frac{1}{\sigma_a}\ \delta_Y(y)\ \sigma_a\ dy$$

$$= \int_{\frac{c-f_a}{\sigma_a}}^{\infty} \sigma_a\ \left(y - \frac{c - f_a}{\sigma_a}\right)\ \delta_Y(y)\ dy$$

$$= \sigma_a\ E\left(\frac{c - f_a}{\sigma_a},\ 0,\ 1\right).$$

D'où:

$$E(c, f_a, \sigma_a) = \sigma_a\ E\left(\frac{c - f_a}{\sigma_a},\ 0,\ 1\right). \qquad (2.3)$$

Il suffit maintenant d'évaluer $E(p, 0, 1)$.

$$E(p, 0, 1) = \int_p^{\infty} (y - p)\ \delta_Y(y)\ dy$$

$$= \int_p^{\infty} (y - p)\ \frac{1}{2\pi}\ e^{-\frac{1}{2}y^2}\ dy$$

$$= \frac{1}{2\pi} \int_p^{\infty} e^{-\frac{1}{2}y^2}\ y\ dy - p \int_p^{\infty} \frac{1}{2\pi}\ e^{-\frac{1}{2}y^2}\ dy$$

$$E(p, 0, 1) = \delta(p) - p\ [1 - F(p)] \qquad (2.4)$$

où $\delta$ et $F$ sont les fonctions de densité et de répartition de la loi normale $N(0, 1)$.

La combinaison de (2.3) et (2.4) implique:

$$E(c, f_a, \sigma_a) = \sigma_a \left( \delta(p) - p\,[1 - F(p)] \right) \qquad (2.5)$$

où $p = \dfrac{c - f_a}{\sigma_a}$

Pour compléter l'étude de $E(c, f_a, \sigma_a)$, voyons quel est le domaine de variation de $f_a$ et comment se calcule $\sigma_a$. Des passagers de plusieurs origines-destinations peuvent utiliser l'avion sur l'arc considéré. Ainsi $f_a$ le nombre de passagers assignés à l'avion se partage en

$$f_a = f_a^1 + f_a^2 + \ldots + f_a^n$$

où $f_a^i$ est le nombre de passagers d'une paire O-D utilisant cet arc. Si $\mu_i$ est le nombre total de passagers d'une de ces O-D, alors $X_a$ la variable aléatoire correspondant au nombre journalier de passagers utilisant l'arc a peut varier entre 0 et $\mu_1 + \mu_2 + \ldots + \mu_n$. L'écart-type de $X_a$ est fonction des $\sigma^i$, les écarts-types des nombres journaliers de passagers de chacune des paires O-D desservies et des $f_a^i$.

D'abord comme $f_a^i$ est une fraction de $\mu_i$, $\sigma_a^i$ l'écart-type du nombre de passagers de la $i^{ième}$ paire O-D utilisant l'arc a se calcule par la relation suivante:

$$\sigma_a^i = \frac{f_a^i}{\mu_i}\,\sigma^i \qquad (2.6)$$

Pour calculer $\sigma_a$ à partir des $\sigma_a^i$ nous utilisons l'hypothèse suivante:

Hypothèse B: Les nombres journaliers de passagers de deux O-D sont si fortement corrélés que l'erreur n'est pas très grande si on

utilise 1 comme coefficient de corrélation.

Cette forte corrélation découle du fait que les jours de pointe des différents marchés utilisant un avion, coïncident généralement. En utilisant l'hypothèse B ont peut calculer le $\sigma_a$ de $f_a^1 + f_a^2$ passagers provenant de deux marchés et ayant respectivement des écarts-types $\sigma_a^1$ et $\sigma_a^2$.

$$\sigma_a = \left((\sigma_a^1)^2 + 2\rho_{12}\,\sigma_a^1\,\sigma_a^2 + (\sigma_a^2)^2\right)^{\frac{1}{2}}$$

où $\rho_{12}$ est le coefficient de corrélation entre $f_a^1$ et $f_a^2$.

$$\sigma_a \approx \left((\sigma_a^1)^2 + 2\sigma_a^1\,\sigma_a^2 + (\sigma_a^2)^2\right)^{\frac{1}{2}} \quad \text{si } \rho_{12} \approx 1$$

$$\sigma_a \approx \sigma_a^1 + \sigma_a^2$$

En étendant cette relation à plusieurs O-D.

$$\sigma_a \approx \sigma_a^1 + \sigma_a^2 + \ldots + \sigma_a^n \qquad (2.7)$$

Combinant (2.6) et (2.7) on obtient:

$$\sigma_a \approx \sum_{i=1}^{n} \frac{f_a^i}{\mu_i}\,\sigma^i \qquad (2.8)$$

Avec ces relations pour calculer le nombre de passagers excédentaires en fonction de $f_a$, il suffit d'estimer les $\mu_i$ et $\sigma^i$.

Les formules (2.4) et (2.5) sont difficilement utilisables sur ordinateur car il n'existe pas d'expression algébrique pour la fonction de répartition de la loi normale. Cette fonction doit être évaluée avec une table. Pour réduire au minimum l'effort de calcul, nous avons développé A(p) une expression algébrique simple qui est

une très bonne approximation de E(p, 0, 1). A(p) est défini comme suit:

$$A(p) = \begin{cases} .4 \times \left(\frac{5 - p}{5}\right)^7 & \text{si } p \leq 5 \\ 0 & \text{si } p > 5 \end{cases} \qquad (2.9)$$

Le tableau 1 illustre par un exemple le comportement de $p = \frac{c - f_a}{\sigma_a}$, $E(p, 0, 1)$, $A(p)$ en fonction de $f_a$ pour un exemple où la capacité $c = 100$ places, le nombre de passagers assignés varie entre 0 et 120 tout en conservant une même proportion des différentes O-D. Dans ce cas $\sigma_a = \lambda \cdot f_a$ avec

$$\lambda = \sum_{i=1}^{n} \frac{f_a^i}{f_a} \cdot \frac{\sigma_i}{\mu_i} = .2$$

Pour $p \in (5, \infty)$ (ce qui correspond à $f_a < 50$) l'erreur en prenant $A(p) = 0$ est négligeable car $0 \leq E(p, 0, 1) \leq .0001$. Pour $p \in [0, 5]$ (ce qui correspond à $f_a \in [50, 100]$), l'erreur est inférieure à .006. Pour $p < 0$ (ce qui correspond à un avion surchargé), nous avons $A(p) > E(p, 0, 1)$. Ainsi l'approximation surestime le nombre de passagers excédentaires quand $f_a > c$ et rend encore plus désavantageuse des solutions qui dépassent la capacité de l'avion. Ceci ne peut qu'accélérer la convergence de l'algorithme vers des solutions acceptables.

TABLEAU I

APPROXIMATION POUR LE CALCUL
DU NOMBRE DE PASSAGERS EXCEDENTAIRES

| $f_a$ | 0 | 20 | 50 | 60 | 70 | 80 | 90 | 100 | 110 | 120 |
|---|---|---|---|---|---|---|---|---|---|---|
| p | ∞ | 20 | 5 | 3.33 | 2.14 | 1.25 | .56 | 0 | -.45 | -.83 |
| E(p,0,1) | .0000 | .0000 | .0001 | .0003 | .006 | .05 | .18 | .40 | .66 | .94 |
| A(p) | 0 | 0 | 0 | .0002 | .008 | .05 | .174 | .40 | .73 | 1.17 |

2.2.3 Calcul du nombre de passagers excédentaires pour une route

D'abord introduisons la notation suivante:

$P_q(h)$ = Espérance du nombre de passagers excédentaires pour la route q. Il s'agit pour une assignation h, du nombre de passagers du marché $r_q$ devant utiliser la route q qui sont refusés faute de place;

$E_a(h) = E(c, f_a, \sigma_a)$ = L'espérance du nombre de passagers refusés sur l'arc a. Il s'agit seulement de simplifier la notation en sous-entendant c et en considérant $f_a$ et $\sigma_a$ comme fonction de h.

Comme un avion est utilisé par des passagers de plusieurs marchés, il faut déterminer à quels marchés appartiennent les passagers excédentaires. L'hypothèse suivante est utilisée:

Hypothèse C: Les passagers assignés à un avion ont une égale probabilité de ne pas pouvoir voyager, indépendamment de leur origine-destination.

Cette hypothèse est réalisée si les places sont attribuées dès que demandées et que les réservations arrivent aléatoirement pour les différentes origines-destinations. Cette hypothèse ne sera pas réalisée si les sièges sont contingentés par origine-destination ou si une priorité est donnée à certaines origines-destinations. Dans de tels cas l'hypothèse C devra être modifiée mais une solution du même type sera probablement encore possible.

Ainsi en supposant que le taux de refus sur un arc (avion) est le même pour toutes les origines-destinations qui l'utilisent, le nombre de passagers d'une route refusés est donné par:

$$\frac{h_q}{f_a} E_a(h) \qquad \text{avec} \qquad f_a = \sum_{q \in E_a} h_q$$

Pour obtenir le nombre total de passagers d'un marché r qui sont refusés le long d'une route, nous posons l'hypothèse suivante:

Hypothèse D: Quand les passagers sont assignés sur les routes de façon optimale, la probabilité est nulle qu'un même passager soit refusé sur plus d'un arc de la route utilisée.

L'utilisation de cette hypothèse présente relativement peu de risque par rapport aux simplifications qu'elle permet. D'abord remarquons que les O-D les plus importantes sont desservies par vols directs; l'hypothèse D n'est pas utilisée pour ces routes d'un seul arc. De plus, dans le type de solutions recherchées par les compagnies aériennes, la probabilité qu'un passager soit refusé sur un arc est faible. Or comme la probabilité d'être refusé sur deux arcs est de l'ordre du carré de cette faible probabilité, elle est pratiquement négligeable. Finalement, en guise de sécurité supplémentaire, les effets de cette hypothèse sont contrôlables en vérifiant dans la solution l'accord entre le nombre de passagers excédentaires et le remplissage des avions. Si pour quelques vols à escales multiples le nombre de passagers excédentaires est trop élevé pour le remplissage proposé, il suffira de considérer que l'avion a quelques places supplémentaires pour rétablir une situation acceptable.

D'autre part au cours de la résolution, avant que la solution optimale ne soit atteinte, il se peut que certaines routes soient surchargées. Dans ce cas, les probabilités d'être refusé sur plusieurs arcs composant une route peuvent être élevées et les termes du deuxième ordre peuvent être non négligeables. Le fait de les négliger augmente le coût associé à ces routes surchargées; ainsi l'algorithme diminuera en priorité le flot sur ces routes surchargées, ce qui accélérera la convergence de l'algorithme.

Avec l'hypothèse D le nombre total de passagers d'un marché r refusés le long d'une route q est donné par la somme des passagers refusés pour chacun des arcs constituant la route.

$$P_q(h) = \sum_{a \varepsilon F_q} \frac{h_q}{f_a} E_a(h)$$

$$P_q(h) = h_q \sum_{a \varepsilon F_q} \frac{E_a(h)}{f_a} \qquad (2.11)$$

## 2.3 Attraction des routes

L'attraction exercée par les différentes routes disponibles sur les passagers d'une origine-destination dépend de plusieurs facteurs: l'heure du départ, les arrêts et correspondances, le type d'avion, etc. Nous ne voulons pas étudier en détail l'influence de ces divers facteurs, influence qui peut varier avec les régions du pays, les saisons, etc. Nous allons nous contenter d'un modèle très simple et laissons l'amélioration de cette partie au soin des utilisateurs éventuels qui disposent d'informations plus complètes sur le comportement des passagers.

### 2.3.1 Facteurs et relations utilisés

Les trois facteurs suivants sont considérés pour déterminer l'attraction d'une route:

a) L'heure de départ

L'heure du départ d'une route s'écarte de l'heure de départ désirée et force le passager qui l'utilise à avancer ou à retarder son départ. Ces écarts ont un effet négatif sur l'attraction.

b) le temps ajouté au voyage

Le temps ajouté au voyage est la différence entre le temps du voyage sur cette route et la durée d'un vol direct. Ce temps mesure l'inconvénient que représente les escales et les correspondances. Pour les paires O-D où il n'y a pas de vol direct, on estime ce temps à l'aide d'une régression linéaire entre les temps et les distances des vols directs existants.

c) autres facteurs

Ce dernier facteur pourrait être ajouté pour tenir compte de l'attraction de certains types d'appareils et des services spéciaux offerts sur certaines routes.

Dans un premier temps, nous considérons que l'attraction est une fonction linéaire de ces facteurs. Le modèle linéaire a été retenu car c'est le plus simple. Si nécessaire, on pourra par la suite modifier cette partie en utilisant des relations plus compliquées sans qu'il soit nécessaire de modifier les autres parties du modèle. Avec un modèle linéaire, l'attraction d'une route est définie par la relation suivante:

attraction = $-K_1$ x (temps en avance du départ) $-K_2$ x (le temps en retard du départ)
$-K_3$ x (temps ajouté au voyage) + (autres facteurs)

où $K_1$, $K_2$, $K_3$ sont des coefficients positifs. (2.12)

Cette relation définit l'attraction des routes pour un passager avec une heure de départ précisée. Comme il n'est pas question de traiter les passagers un à un, il faut agréger l'attraction pour le groupe de passagers d'un marché. Le coefficient d'attraction d'une route pour les passagers d'un marché est défini en prenant la moyenne des attractions pondérées par le nombre de passagers voulant partir à chaque heure.

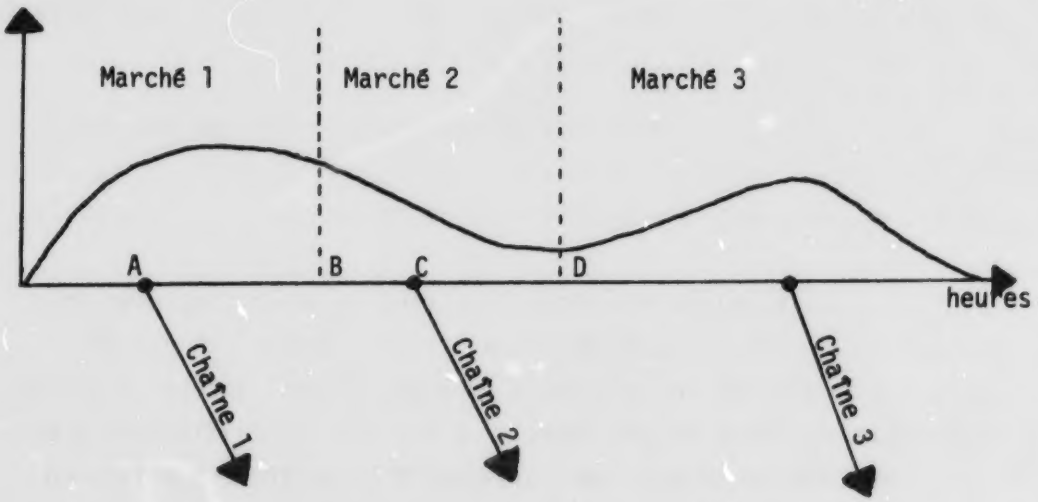
Ce choix de négliger le détail des heures de départ souhaitées se justifie par les simplifications qu'il permet et le fait qu'il est assez utopique de tenter d'être plus précis sur ce point. En effet, la relation définissant l'attraction n'est qu'une première approximation et les estimations d'heures de départ souhaitées sont très approximatives.

### 2.3.2 Regroupement des passagers par heures de départ

Comme nous l'avons vu lors de la définition du réseau, l'ensemble des passagers regroupés dans un marché est défini comme ceux pour qui la route associée est la plus attrayante. Ces groupes de passagers correspondent à des intervalles d'heures de départ tel que l'illustre la figure 6.

FIGURE 6

PARTAGE EN MARCHÉ DES PASSAGERS D'UNE O-D



Ainsi, l'intervalle d'heure de B à D correspond aux heures de départ pour lesquels la chaîne 2 est la plus attrayante. Il est important de rappeler, pour éviter toutes fausses interprétations, que les passagers d'un marché ne voyageront pas tous par la chaîne associée. Ces passagers peuvent aussi voyager par les deux routes formées en utilisant les chaînes adjacentes quand il n'y a pas suffisamment de places sur la route formée avec la chaîne associée.

La frontière entre le marché 1 et le marché 2 est placée au point B d'égale attraction entre la route 1 et la route 2. Ainsi, comme l'attraction d'une route augmente à mesure que l'on s'approche de l'heure de départ de cette route, l'attraction de la route 1 sera plus grande pour les heures de départ de l'intervalle (A, B) et l'attraction de la route 2 sera plus grande pour les heures de départ de l'intervalle (B, C).

Finalement, il faut remarquer que le temps forme un cycle de 24 heures et ainsi le premier vol suit le dernier. Il y a donc une frontière à déterminer entre ces deux vols et ce n'est pas nécessairement la fin de la journée.

## 2.4 Modèle P1: Minimisation de la perte de revenus résultant des passagers excédentaires et de l'utilisation de routes non attrayantes

Rappelons que le problème est de déterminer comment le flot de passagers se déplace des origines aux destinations en tenant compte de l'attraction des routes, des heures de vols et de la limite de capacité des avions. Le modèle P1 consiste à répartir les passagers de chaque marché entre les routes de façon à maximiser le revenu associé aux

passagers et minimiser l'insatisfaction découlant de l'utilisation de routes non attrayantes. Le revenu est la différence entre le revenu total de la vente des billets à la clientèle estimée moins l'espérance du revenu perdu, faute de place. Comme le revenu total des billets est une constante, nous traiterons un problème équivalent de minimiser les revenus perdus, faute de place, plus le coût associé à l'utilisation de routes non attrayantes.

Définissons:

$Z_1$ = Espérance total du revenu perdu en refusant des passagers, faute de place

$Z_2$ = Coût associé à l'utilisation des routes non attrayantes

Alors le problème se résume comme suit:

Minimiser $Z_1 + KZ_2$.

s.a. la demande pour chaque marché est répartie entre les routes assurant la liaison.

Dans cette section, on développe les équations reliant $Z_1$ et $Z_2$ aux flots sur les routes. L'évaluation de la constante de calibration sera discutée en même temps que l'expérimentation au chapitre 4.

### 2.4.1 Perte de revenus résultant des passagers excédentaires

Nous avons établi en 2.2.4 la relation (2.11) qui donne l'espérance du nombre de passagers excédentaires pour une route q.

$$P_q(h) = h_q \sum_{a \in F_q} \frac{E_a(h)}{f_a} \qquad (2.11)$$

Le revenu perdu pour les passagers de cette route, refusés, faute de place, est obtenu en multipliant par $B^r$ le coût du billet de l'origine-destination desservi par la route q;

$$B^r h_q \sum_{a \in F_q} \frac{E_a(h)}{f_a}$$

Le revenu perdu pour l'ensemble du réseau est obtenu en prenant la somme sur toutes les routes desservant tous les O-D.

$$Z_1 = \sum_r \sum_{q \in Q_r} B^r h_q \sum_{a \in F_q} \frac{E_a(h)}{f_a} \qquad (2.13)$$

2.4.2 Coût associé à l'utilisation de routes non attrayantes

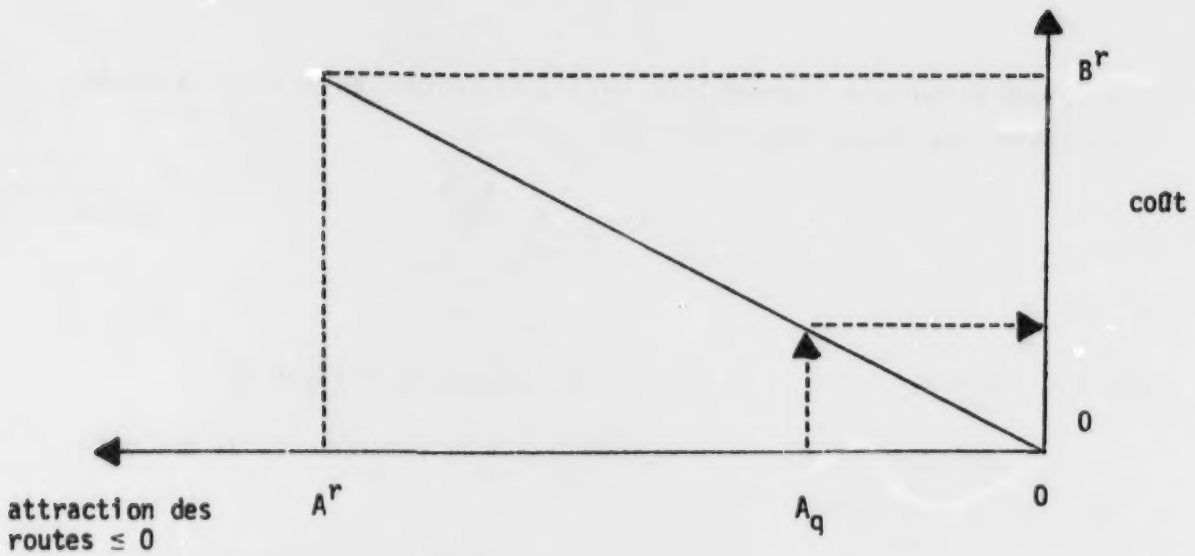
Nous avons établi à la section 2.3.1 la relation (2.12) qui définit l'attraction d'une route q.

$$A_q = -\sum_{i=1}^{3} K_i \cdot t_i \quad \text{avec} \quad \begin{array}{l} t_1 = \text{temps d'attente positif} \\ t_2 = \text{temps d'attente négatif} \\ t_3 = \text{temps ajouté au voyage} \\ K_i \geq 0 \end{array}$$

$|A_q|$ peut être interprété comme le niveau d'insatisfaction de la route q. Pour associer un coût à cette insatisfaction, considérons $A^r$ le niveau d'attraction minimum au delà duquel les passagers du marché r ne l'utilisent plus. Nous avons considéré que ce niveau augmente linéairement avec la distance de l'origine à la destination. Définissons $C^r$ le facteur de conversion de l'insatisfaction en coût tel que $C^r|A^r| = B^r$. Ainsi, le coût par passager associé à l'utilisation de la route q est $C^r|A_q|$. La figure suivante illustre les propriétés de ce facteur de conversion. Le coût est nul pour une route idéale ayant une attraction égale à zéro. Le coût augmente linéairement à mesure que la route devient moins attrayante pour devenir égal au coût du billet quand l'attraction descend au minimum acceptable.

FIGURE 7

COUT ASSOCIE A L'INATTRACTION DES ROUTES




Le coût total associé à l'insatisfaction de tous les passagers utilisant la route q est donné par

$$h_q \, C^r \, |A_q| = - h_q \, C^r \, A_q$$

Le coût total associé à l'inattraction des routes est pour le marché r:

$$-C^r \sum_{q \in Q_r} h_q \, A_q$$

Le coût pour tout le réseau est:

$$Z_2 = - \sum_r C^r \sum_{q \in Q_r} h_q \, A_q \qquad (2.14)$$

Cette définition de $Z_2$ calibre relativement bien les coûts de l'inattraction des routes de différents marchés l'un par rapport à l'autre. Par contre, la définition de K qui calibre les coûts de l'inattraction relativement aux coûts des passagers refusés faute de place, sera discutée au chapitre 4.

### 2.4.3 Unicité et stabilité de la solution

Les algorithmes de résolution trouvent des optimums locaux; la convexité de l'objectif assurerait d'atteindre un optimum global. Comme l'objectif n'est pas toujours convexe, la solution trouvée peut être optimale mais rien ne le garantit. Nous présentons en détail à l'annexe 2 un cas où la fonction économique est non-convexe et discutons les situations particulières la rendant convexe. Résumons ici les principaux résultats.

C'est la non convexité du terme $Z_1$ qui cause celle de la fonction économique, le terme $Z_2$ étant linéaire. Rappelons la définition de $Z_1$

$$Z_1 = \sum_r \sum_{q \in Q_r} B^r h_q \sum_{a \in F_q} \frac{E_a(h)}{f_a}$$

$$E_a(h) = \sigma_a E(p,0,1)$$

$$\sigma_a = \sum_{q \in F_a} h_q \sigma_{r_q} / u_{r_q}$$

$$E(p,0,1) = f(p) - p[1-F(p)]$$

Les principaux résultats démontrés en annexe sont:

$1^o$) $E(p,0,1)$ est convexe.

$2^o$) Un exemple montre que $E_a(h)$ peut être non convexe.

3°) $E_a(h)$ est convexe si les coefficients de variations de tous les marchés sont égaux.

4°) $E_a(h)/f_a$ est convexe pour $f_a < c$ si les coefficients de variations de tous les marchés sont égaux et inférieurs à .4.

5°) Un exemple montre que $Z_1$ peut être non convexe même quand $E_a(h)/f_a$ est convexe.

6°) $Z_1$ est convexe si les coefficients de variations de tous les marchés sont égaux et si les coûts des passagers refusés ne dépendent pas de l'O-D mais seulement des arcs où ils sont refusés.

En pratique, probablement que $Z_1$ est convexe car les coûts des billets des passagers utilisant un même arc ne sont pas très différents, sûrement pas aussi différents que dans l'exemple non-convexe présenté en annexe. De même les coefficients de variations ne sont pas très différents, ils sont surtout concentrés entre .3 et .4. Ainsi, une somme de termes dont la très grande majorité sont convexes, devrait être convexe.

Discutons maintenant la stabilité de la solution. La propriété recherchée est la stabilité de la solution pour de faibles variations des données d'entrée. Cette qualité est importante, car les données sont toujours approximatives et il faut une solution s'approchant le plus possible de l'optimum.

Dans certains cas,cependant, de faibles erreurs sur l'attraction des routes peuvent entraîner selon ce modèle, de grandes variations de la répartition des passagers. Dans le cas où les avions desservant une paire O-D sont peu remplies, le terme non linéaire mesurant la perte de revenus résultant des passagers excédentaires est petit et la répartition des passagers dépend surtout du terme linéaire mesurant l'attraction des routes. Ainsi, nous nous retrouvons avec un problème presque

linéaire. Il est bien connu que la solution d'un problème d'optimisation linéaire est extrémiste et instable. Dans ce cas, la solution consistera à placer tous les passagers de cette O-D sur la route la plus attrayante et ainsi,minimiser le coût associé à l'utilisation de routes moins attrayantes.

Dans le cas où il existe une deuxième route presqu'aussi attrayante que la première, il suffit d'une petite erreur d'estimation de l'attraction pour que tous les passagers soient déplacés sur la deuxième route.

Même s'il y a suffisamment de passagers pour que plus d'une route soit utilisée, la solution reste relativement extrémiste et instable tant que le coefficient moyen de remplissage reste inférieur à 60%. Pour illustrer ce comportement, voyons un exemple de la répartition des passagers d'une paire O-D exprimé en fonction du nombre total de passagers.

Considérons une paire O-D avec 3 routes constituées d'un seul arc et toutes desservies par des avions de 100 places. Prenons le coût du billet égal à $100. et les coûts mesurant l'inattraction des routes respectivement égaux à $10, $15, $50.

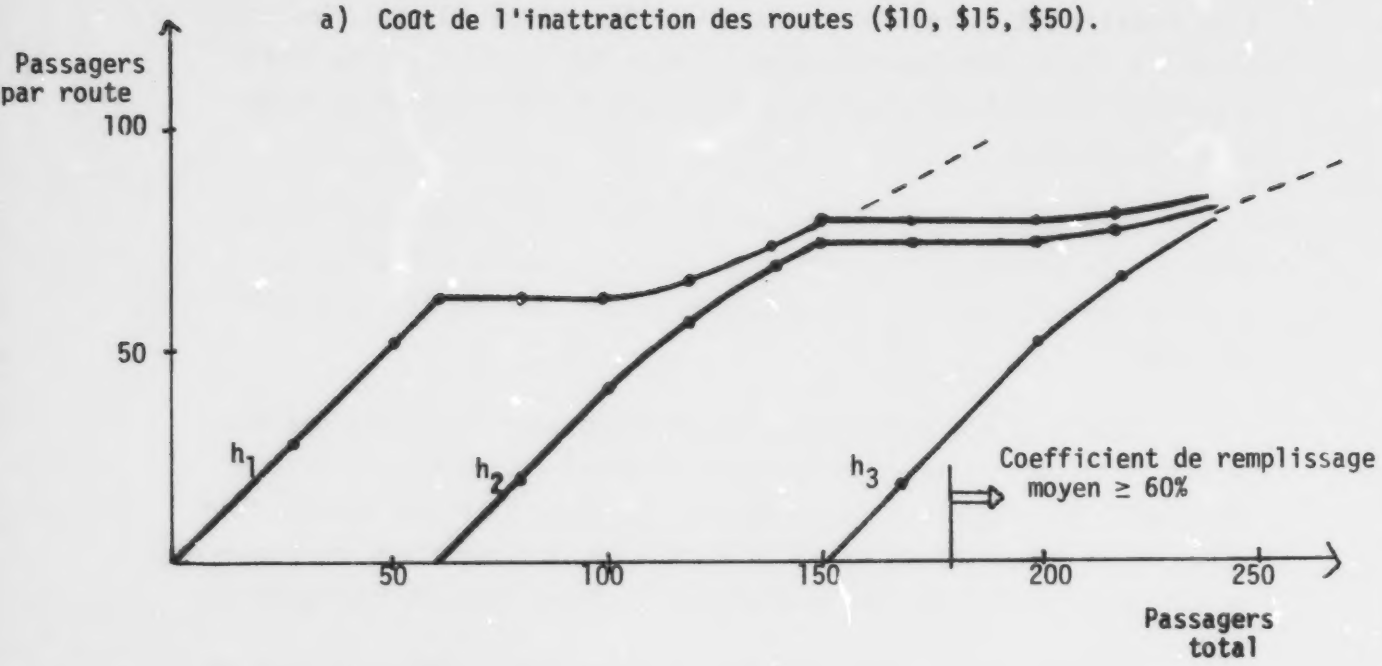
La figure 8a présente sur un même graphique les nombres de passagers utilisant chacune des trois routes en fonction du nombre de passagers total. Cet exemple présente une répartition extrémiste pour un nombre de passagers inférieur à 170. Par contre, la répartition devient plus équilibrée par la suite quand le coefficient moyen de remplissage dépasse 60%. La répartition est instable quand le coefficient moyen de remplissage est faible. Si les coûts associés à l'inattraction des routes 1 et 2 sont remplacés par $14, $13 la répartition des passagers est complètement changée comme nous pouvons le voir à la figure 8b.

En conclusion, ce modèle est satisfaisant pour les coefficients de remplissage élevés, mais présente des faiblesses pour les faibles coefficients de remplissage. Dans les prochaines sections, nous allons palier à ce problème en présentant des modèles équilibrés et stables même pour les faibles remplissages.

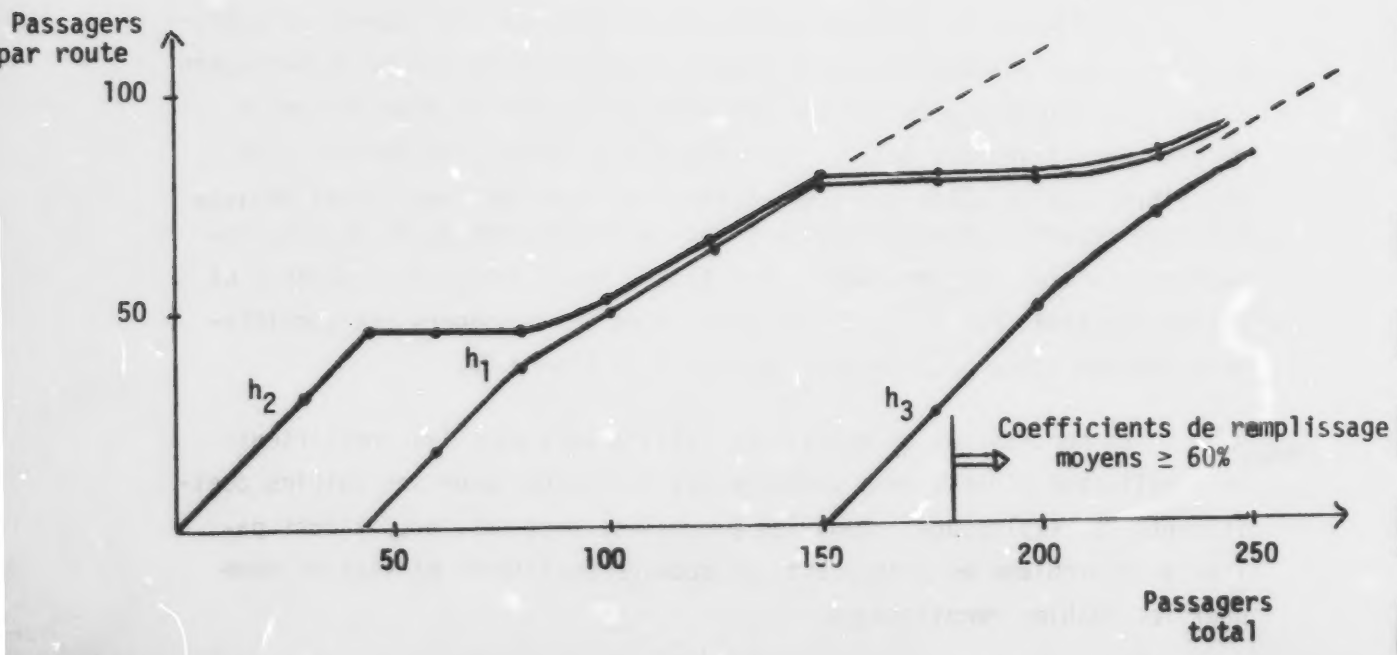
Figure 8

Exemple de répartition des passagers entre 3 routes en fonction du nombre total de passagers.

a) Coût de l'inattraction des routes ($10, $15, $50).




b) Coût de l'inattraction des routes ($14, $13, $50).



## 2.5 Modèle P2: Compromis entre la perte de revenu due aux passagers excédentaires et une répartition en proportion de l'attraction des routes.

Nous développons maintenant un deuxième modèle moins extrémiste et plus stable que le modèle précédent. Ce second modèle réalise aussi une "optimisation coopérative". L'objectif global à minimiser est encore composé des deux mêmes facteurs. Le revenu perdu, faute de place, et un coût associé au choix de routes moins attrayantes. Ce modèle diffère du précédent par la façon d'exprimer le deuxième facteur. Le terme linéaire est remplacé par un terme non-linéaire et cette modification rend le modèle moins extrémiste.

La nouvelle formulation repose sur une nouvelle modélisation de l'attraction des routes. Dans la prochaine section, on discute la différence entre ces deux notions d'attraction en considérant les hypothèses correspondantes de comportement des passagers.

### 2.5.1 Comportement des passagers en fonction de l'attraction des routes

Dans le premier modèle en minimisant la somme des insatisfactions, on obtient un cheminement de passagers correspondant à l'hypothèse de comportement suivante:

> Chaque passager utilise la route la plus attrayante qui comporte encore des sièges disponibles.

Suivant ce comportement, si les avions n'avaient pas de limite de capacité, tous les passagers d'un marché utiliseraient la route la plus attrayante. Avec la limite de capacité, les passagers saturent les routes dans l'ordre de leur attraction. Ce comportement est illustré par l'exemple présenté à la figure 8.

Un comportement aussi monolithique des passagers ne correspond pas à la réalité. Les passagers ne sont généralement pas unanimes sur

l'ordre de préférence des routes. Certains sont sensibles aux nombres d'escales et de correspondances, d'autres attachent plus d'importance à l'heure de départ ou à l'heure d'arrivée du vol. Ainsi, même sans que la limite de capacité des avions intervienne, les passagers se répartissent entre plusieurs routes à cause des différences de préférences individuelles.

Pour représenter ce comportement dans le second modèle, on définit le coefficient d'attraction d'une route comme étant la proportion des passagers qui préfèrent l'utiliser. Ces coefficients peuvent être obtenus par voie de sondage . On pourrait aussi développer une fonction qui indiquerait la répartition des préférences suivant l'horaire, les escales et les correspondances pour les routes disponibles. L'hypothèse de comportement associée à ce second modèle est la suivante:

> Les passagers se répartissent entre les routes en proportion des coefficients d'attraction en autant que le permet la capacité des avions.

Ce comportement est obtenu en modifiant en conséquence la fonction objective.

### 2.5.2 Evaluation des coefficients d'attraction

Les coefficients d'attraction doivent être définis pour chaque groupe de passagers voulant partir à une heure donnée. Il faut déterminer la proportion des passagers préférant utiliser chacune des trois routes accessibles. Il convient de préciser ici que l'heure de départ désirée est l'heure à laquelle le passager souhaiterait avoir un vol disponible. Ce sont les coefficients d'attraction qui doivent indiquer que les passagers se réajustent en fonction des routes offertes.

L'information de base sur la distribution des heures de départ souhaitées et les coefficients d'attraction des routes pour les groupes de passagers voulant partir à une heure donnée peuvent être obtenus par un sondage qui peut probablement être réalisé par le système de réservation.

A long terme, pour diminuer la cueillette de l'information nécessaire à l'estimation des coefficients d'attraction, il serait possible de développer un modèle pour les déterminer, c'est-à-dire qu'on pourrait exprimer l'attraction des 3 routes en fonction du type de marché, de l'horaire des escales et correspondances de ces routes. On pourrait calibrer et valider ce modèle avec l'information obtenue lors des sondages précédents. Ce modèle aurait l'avantage de donner des coefficients d'attraction pour n'importe quel horaire. Cette souplesse est très importante pour le développement d'un modèle automatisé de fabrication d'horaires.

### 2.5.3 Objectifs correspondant à ce comportement des passagers.

L'objectif doit conduire à une répartition des passagers réalisant un compromis entre le respect de la limite de capacité des avions et la répartition en proportion de l'attraction des routes. Cet objectif sera ici encore de la forme $Z_1 + Z_2$, ou $Z_1$ est la perte de revenu, associée aux passagers excédentaires, définie précédemment et $Z_2$ une nouvelle mesure du coût associée à l'insatisfaction des passagers si on s'écarte de leurs préférences. Une fonction $Z_2$ satisfaisante peut être définie de la façon suivante:

soit $A_q \geq 0$ le coefficient d'attraction de la route q,

$\left(\sum_{q \in Q_r} A_q = 1\right)$, soit $d^r$ la demande totale pour le marché r, et

soit g fonction monotone croissante.

$$\text{Soit } Z_2(h) = \sum_r \left\{ \sum_{\substack{q \in Q_r \\ A_q \neq 0}} \int_0^{h_q} g\left(\frac{x}{A_q}\right) dx + \sum_{\substack{q \in Q_r \\ A_q = 0}} g(d^r)\, h_q \right\}$$

Etudions d'abord les propriétés de cet objectif en l'absence de limite de capacité des avions. La proposition suivante montre que dans ces conditions l'objectif $Z_2$ donne une répartition en proportion des coefficients d'attraction.

Proposition: La solution $h^*$ du problème suivant (2.15)

$$\text{Min } Z_2(h) \text{ s.a. } \sum_{q \in Q_r} h_q = d^r \text{ pour tout } r \text{ et } h_q \geq 0 \text{ pour tout } q$$

est définie par

$$h_q^* = d^r \; A_q \text{ pour tout } r \text{ et tout } q \in Q_r$$

Preuve:

$$\frac{\partial Z_2(h)}{\partial h_q} = \begin{cases} g(h_q / A_q) & \text{si } A_q > 0 \\ g(d^r) & \text{si } A_q = 0 \end{cases}$$

Comme g est croissante, $Z_2$ est strictement convexe et par conséquent $h_q^*$ est optimal si et seulement si les conditions de Kuhn-Tucker sont vérifiées.

$(\forall r)\ (\exists u_r)$ tel que $(\forall q \in Q_r)$

$$\frac{\partial Z_2(h)}{\partial h_q} \geq u_r \quad \text{si} \quad h_q = 0 \tag{2.16 a}$$

$$\frac{\partial Z_2(h)}{\partial h_q} = u_r \quad \text{si} \quad h_q > 0 \tag{2.16 b}$$

$$\sum_{q \in Q_r} h_q = d^r \tag{2.16 c}$$

$$h_q \geq 0 \tag{2.16 d}$$

Montrons que $h_q^*$ satisfait les conditions 2.16

d) $h_q^* \geq 0$ car $h_q^* = d^r A_q$ , $d^r > 0$ et $A_q \geq 0$

c) $\sum_{q \in Q_r} h_q^* = d^r \sum_{q \in Q_r} A_q = d^r$ .

b) Si $h_q^* > 0$ , alors $A_q > 0$ et $h_q^* / A_q = d^r$ .

Ainsi $g\left(\frac{h_q^*}{A_q}\right) = g(d^r)$ ce qui implique que $\frac{\partial Z_2(h^*)}{\partial h_q} = g(d^r)$

En prenant $u_r = g(d^r)$ la condition est satisfaite

a) Si $h_q^* = 0$, alors $A_q = 0$ et $\frac{\partial Z_2(h^*)}{\partial h_q} = g(d^r)$

et donc $\frac{\partial Z_2(h^*)}{\partial h_q} \geq u_r$

La sous-famille de fonction obtenue en prenant $g(y) = \alpha (y-\beta)^\gamma$, $\gamma > 0$ permet par exemple de définir un éventail de coût avec des taux de croissance plus ou moins rapide suivant le choix des paramètres.

Finalement en prenant comme objectif $Z_1(h) + Z_2(h)$, nous obtenons une répartition tenant compte du coût de dépasser la capacité des avions et du coût de s'écarter des préférences des passagers.

Tout comme pour le modèle précédent, l'objectif n'est pas nécessairement convexe à cause du terme $Z_1$. Par contre, la présence d'un deuième terme convexe plutôt que linéaire devrait agrandir le domaine où l'objectif est convexe.

2.5.4 Stabilité de la solution

Pour montrer comment cette modification a amélioré la stabilité de la solution, nous allons reprendre avec ce modèle l'exemple de la section 2.4.3. Pour cet exemple nous prenons une variante des plus simples du modèle en utilisant $g(y) = (1/5)\, y$. Le facteur d'échelle est nécessaire pour pondérer le terme $Z_2$ en accord avec le coût du billet de l'exemple. La fonction $Z_2$ devient

$$Z_2(h) = \sum_r \sum_{\substack{q \in Q_r \\ A_q \neq 0}} \frac{h_q^2}{10\, A_q}$$

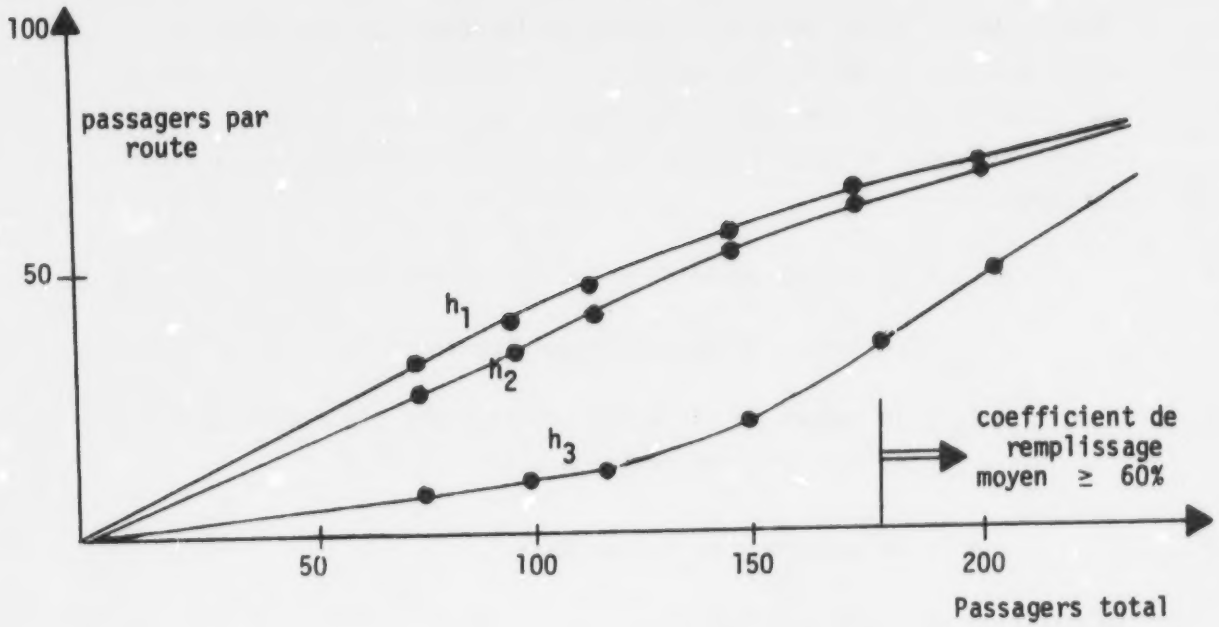
Les coûts associés à la non attraction des routes étaient respectivement \$10, \$15, \$50. Supposons que les proportions des passagers qui désirent utiliser chacune de ces trois routes sont $A_1 = .5$, $A_2 = .4$, $A_3 = .1$ . Ces proportions sont compatibles avec l'attraction des routes.

La figure 9 présente sur un même graphique le nombre de passagers utilisant chacune des 3 routes en fonction du nombre de passagers total. Pour les taux moyens d'occupation supérieurs à 60% ces nombres sont semblables à ceux obtenus avec le modèle P1. Par contre pour les plus faibles taux d'occupation, les courbes associées au modèle P2 sont nettement plus régulières.

Il est facile de constater que les solutions de ce modèle sont plus stables. Même si l'attraction des deux premières routes était interchangée, les changements de la répartition des passagers ne seraient par trop importants. Les courbes de $h_1$ et $h_2$ seraient interchangées et ainsi, une variation de 20% des coefficients d'attraction changerait les

FIGURE 9

EXEMPLE DE REPARTITION DES PASSAGERS ENTRE 3 ROUTES EN FONCTION DU NOMBRE TOTAL DE PASSAGERS




répartitions des passagers d'au plus 20%, et de beaucoup moins pour les taux d'occupation élevés.

## 2.6 Modèle P3: Répartition des passagers en considérant l'"optimisation égoïste"

La différence entre l'"optimisation égoïste" et l'"optimisation coopérative" a été présentée à la section 1.2.2. Rappelons que le comportement des passagers suivant l'"optimisation égoïste" est le suivant:

chaque passager choisit son itinéraire selon ses préférences tout en devant se contenter de seconds choix si l'avion qu'il veut utiliser est déjà rempli par d'autres passagers.

Contrairement aux modèles précédents qui déterminaient une assignation des passagers réalisant un équilibre global entre l'attraction et la capacité, ce modèle reproduit le détail du processus du choix des routes par les passagers. Il distingue entre les passagers qui essaient d'utiliser une route et ceux qui y trouvent place et l'utilisent en fait. A cette fin, les notations sont modifiées de la façon suivante:

$a_q$ = le nombre de passagers qui essaient d'utiliser la route q

$k_q$ = le nombre de passagers qui trouvent place sur la route q

$P_q$ = le nombre de passagers qui essaient d'utiliser la route q et sont refusés, faute de place

$d^r$ = la demande pour le marché r

La relation suivante est immédiate:

$$a_q = k_q + P_q \tag{2.17}$$

Comme un passager peut essayer d'utiliser la route $q_1$, y être refusé, faute de place et ensuite essayer d'utiliser la route $q_2$, il peut être compté dans $a_{q_1}$ et $a_{q_2}$. Ainsi nous avons:

$$\sum_{q \in Q_r} a_q \geq d^r$$

D'autre part comme chaque passager voyage sur une et une seule route, nous avons la relation:

$$\sum_{q \in Q_r} k_q = d^r \tag{2.18}$$

Ce modèle considère que la demande est fixe et que tous les passagers peuvent voyager. Le prochain modèle traitera le cas d'une demande variant en fonction de l'attraction et de la disponibilité des routes offertes.

2.6.1 Comportement des passagers

L'hypothèse de comportement des passagers face à l'attraction des routes est celle du modèle précédent.

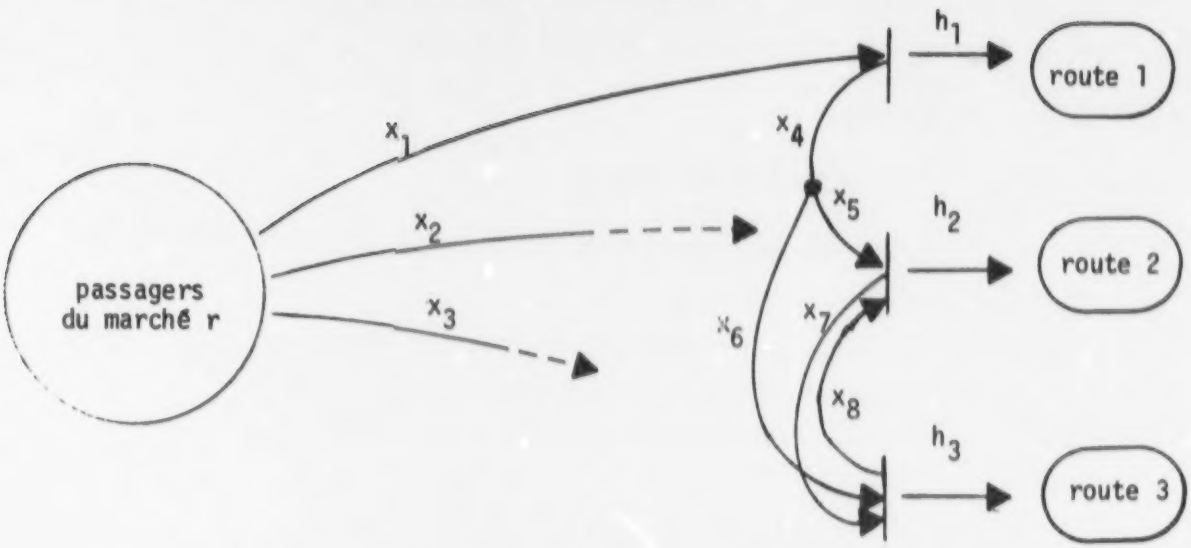
> " Les passagers tentent de se répartir entre les routes en proportion des coefficients d'attraction ($A_q$) en autant que le permet la capacité des avions".

Ce qui distingue ce modèle du précédent, c'est qu'il considère de plus une hypothèse de comportement des passagers quand ils sont refusés, faute de place dans un avion. L'hypothèse doit décrire comment les passagers font un deuxième choix et même un troisième choix quand ils ne peuvent utiliser les routes choisies. La figure 10 présente la complexité des choix possibles. Certains arcs ont été supprimés pour des raisons évidentes.

Cette figure illustre un comportement où les passagers du marché r tentent d'abord de se répartir entre les 3 routes en proportion des coefficients d'attraction ce qui correspond aux arcs $x_1$, $x_2$, $x_3$. L'arc $x_4$ représente les passagers refusés, faute de place, sur la route 1, ces passagers se répartissent en $x_5$ et $x_6$ suivant les coefficients d'attraction des routes 2 et 3. Une partie des passagers de l'arc $x_5$ peut être refusée, faute de place, sur la route 2, et dans ce cas ils utilisent la route 3 comme l'indique l'arc $x_7$. De même l'arc $x_8$ correspond à des passagers qui en sont à leur troisième choix. Ici nous travaillons sous l'hypothèse que tous les passagers trouvent finalement place. Dans le cas général, certains passagers des arcs $x_7$ et $x_8$ pourraient être refusés, faute

FIGURE 10

CHEMINEMENT DES PASSAGERS




de place, et ne pas pouvoir voyager. Les $k_i$ comprennent les passagers de toute provenance qui finalement utilisent la route i. L'hypothèse de comportement des passagers refusés faute de place, qui correspond à la figure 10 est la suivante:

> Les passagers ayant été refusés, faute de place, se répartissent entre les routes non essayées en proportion des coefficients d'attraction.

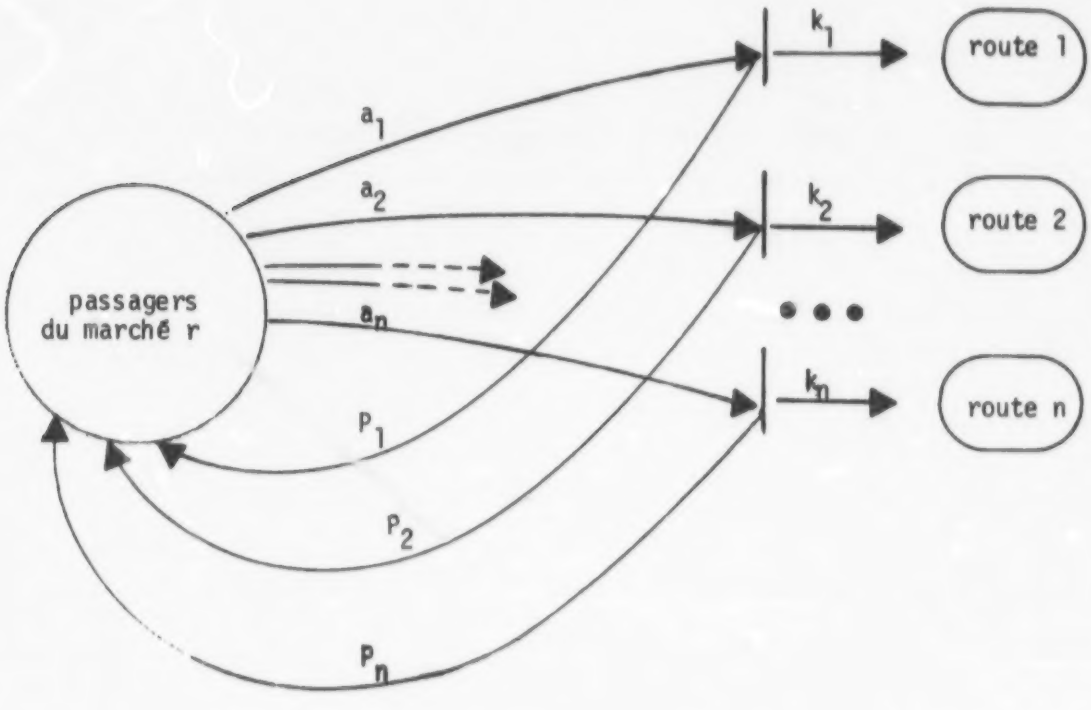
Comme il est facile de le supposer à la vue de la figure 10, cette hypothèse de comportement conduit à un système d'équations très complexe. Nous proposons d'utiliser plutôt une hypothèse de comportement légèrement différente qui donne lieu à un modèle beaucoup plus simple. Cette variante que nous appellerons l'hypothèse oubliante se formule comme suit:

> Les passagers ayant été refusés, faute de place, se répartissent en proportion des coefficients d'attraction entre toutes les routes qu'elles aient été essayées précédemment ou pas.

Ainsi les passagers refusés font un autre essai en oubliant où ils ont été refusés. Ils se comportent donc au deuxième et troisième essai exactement comme au premier. Ce comportement est représenté à la figure 11.

FIGURE 11

CHEMINEMENT DES PASSAGERS SUIVANT L'HYPOTHESE OUBLIANTE




Les $a_i$, $k_i$, $P_i$ correspondent aux définitions du début de la section 2.6. Ainsi $a_i$ est le nombre total de passagers qui essaient d'utiliser la route i qu'ils en soient à leur premier, deuxième ou $n^{ième}$ essai. De même $k_i$ est le nombre de passagers qui finalement utilisent la route i

quelque soit le cheminement qu'ils aient fait avant. Les $R_i$ sont aussi des nombres totaux de passagers refusés.

L'hypothèse oubliante peut permettre à un passager d'essayer à plusieurs reprises d'utiliser la même route. Ce cas est rare car la probabilité d'être réfusé est rarement supérieure à 5% et seulement la partie $A_i$ de ces passagers réessaie la route i. En pratique aussi une partie des passagers refusés réessaie la même route quand ils déplacent leurs dates de départ pour obtenir la route désirée.

### 2.6.2 Equation du modèle

En combinant l'hypothèse de répartition en proportion des coefficients d'attraction et l'hypothèse oubliante, nous obtenons que les $a_i$ sont proportionnels aux $A_i$. En effet s'il s'agit d'un premier essai, l'hypothèse de proportionalité s'applique, s'il s'agit d'un essai subséquent, l'hypothèse oubliante assure la même proportionalité. Ainsi il doit exister un $u_r$ tel que $a_q = u_r A_q$ pour tout $q \in Q_r$. En substituant (2.17) il existe $u_r$ tel que

$$k_q + P_q = u_r A_q \text{ pour tout } q \in Q_r \qquad (2.19)$$

Le nombre de passagers excédentaires sur la route q a été défini par l'équation (2.11)

$$P_q = h_q \sum_{a \in F_q} \frac{E_a(h)}{f_a} \text{ avec } f_a = \sum_{q \in E_a} h_q$$

Il faut se rappeler que h correspond aux nombres totaux de passagers essayant d'utiliser chaque route et non aux nombres de passagers qui y trouvent place. A l'annexe 3 on trouve une façon d'exprimer $P_q$ comme

fonction des nombres de passagers $k_q$ qui utilisent chacune des routes q. Ainsi, nous obtenons un système d'équations qui avec (2.18) caractérise le $k_q$

Pour tout r il existe un $u_r$ tel que:

$$k_q + P_q(k) = u_r A_q \qquad (\forall q \in Q_r)$$

$$\sum_{q \in Q_r} k_q = d^r \tag{2.20}$$

Comme par définition $k_q$, $P_q$, $A_q \geq 0$, quand $A_q = 0$ il découle immédiatement de (2.19) que $k_q = 0$. Ainsi, ces routes peuvent être éliminées du réseau car elles ne sont pas utilisées. Par la suite, nous considérons seulement les routes ayant $A_q > 0$.

Plutôt que de résoudre directement le système (2.20), transformons-le pour retrouver la structure des modèles précédents. Ceci permet de le résoudre avec le même algorithme.

Comme le cas $A_q = 0$ est exclus, il est immédiat que le système (2.20) est équivalent au système suivant.

Pour tout r il existe $u_r$ tel que

$$\frac{k_q + P_q(k)}{A_q} \begin{cases} = u_r & \text{si } k_q > 0 \\ \geq u_r & \text{si } k_q = 0 \end{cases} \quad (\forall q \in Q_r)$$

$$\sum_{q \in Q_r} k_q = d^r$$

$$k_q \geq 0 \qquad (\forall q \in Q_r) \tag{2.21}$$

Ce système a la même structure que celle des conditions de Kuhn-Tucker des problèmes d'optimisation associés aux modèles précédents. Rappelons que ces problèmes avaient la structure suivante:

$$\text{Min} \quad f(h)$$

$$\sum_{q \in Q_r} h_q = d^r \qquad (\forall r)$$

$$h_q \geq 0 \quad (\forall q \in Q_r) \ (\forall r) \tag{2.22}$$

Les conditions de Kuhn-Tucker pour ces problèmes sont donc:
pour tout r il existe un $u_r$ tel que

$$\frac{\partial f(h)}{\partial h_q} \quad \left\{ \begin{array}{ll} = u_r & \text{si} \quad h_q > 0 \\ \geq u_r & \text{si} \quad h_q = 0 \end{array} \right\} \quad (\forall q \in Q_r)$$

$$\sum_{q \in Q_r} h_q = d^r$$

$$h_q \geq 0 \qquad (\forall q \in Q_r) \tag{2.23}$$

Les systèmes (2.21) et (2.23) présentent la même structure en faisant correspondre $k_q$ à $h_q$ et $\frac{k_q + P_q(k)}{A_q}$ à $\frac{\partial f(h)}{\partial h_q}$ .

Cette observation suscite la question de l'existence d'un objectif f tel que les conditions (2.21) soient les conditions de Kuhn-Tucker d'un problème d'optimisation. Malheureusement il n'existe pas de tel objectif car la condition suivante n'est pas satisfaite:

$$\frac{\partial}{\partial k_p}\left(\frac{k_q + P_q(k)}{A_q}\right) = \frac{\partial}{\partial k_q}\left(\frac{k_p + P_p(k)}{A_p}\right) \tag{2.24}$$

Cette condition découle du résultat d'analyse bien connu suivant: pour toute fonction continue f

$$\frac{\partial f(h)}{\partial h_p \, \partial h_q} = \frac{\partial f(h)}{\partial h_q \, \partial h_p}$$

**Par contre, la condition (2.24) est vraie si p et q sont deux** routes desservant le même marché r. En effet dans ce cas comme les deux routes n'ont pas d'arc en commun et

$$\frac{\partial}{\partial h_p}\left(\frac{h_q - P_q(h)}{A_q}\right) = 0 = \frac{\partial}{\partial h_q}\left(\frac{h_p - P_p(h)}{A_p}\right) \qquad (2.25)$$

Ainsi, chaque sous-problème correspondant à un marché peut être traité comme un problème d'optimisation car il existe un objectif associé. Ces sous-problèmes peuvent donc être traités avec le même algorithme que celui pour les sous-problèmes des modèles précédents.

Même si le problème global défini par les équations (2.21) ne peut pas être ramené à un problème d'optimisation, il possède quand même une structure intéressante. C'est un problème de complémentarité non-linéaire. Nous utiliserons des résultats de la théorie de la complémentarité pour montrer l'existence de solution du problème et pour montrer qu'il peut se résoudre par décomposition en sous-problèmes correspondant aux marchés. Ces développements sont présentés au chapitre 3 qui traite des méthodes de résolutions.

### 2.6.3 Propriété du modèle

D'abord la solution est stable et équilibrée comme celle du **modèle précédent, contrairement à la répartition extrémiste provenant des modèles linéaires.**

De plus, ce modèle n'a pas besoin de calibration. L'assignation des passagers est complètement déterminée par les hypothèses de comportement. Il n'est donc pas nécessaire d'ajuster de facteur de pondération entre l'influence de l'attraction des routes et celle de la limite de capacité des avions.

Finalement pour mieux comprendre l'approximation associée à l'hypothèse oubliante, exprimons $k_q$ (le nombre de passagers utilisant la route $q \in Q_r$) en fonction de la demande, des coefficients d'attractions et des nombres de passagers en excès pour les différentes routes desservant le même marché. Comme seules les routes avec l'attraction positive sont considérées ($A_q > 0$), l'équation (2.19) pour la route q permet d'écrire

$$\frac{k_q + P_q}{A_q} = u_r \tag{2.26}$$

Pour chacune des autres routes $p \in Q_r$ l'équation (2.19) associée devient quand on y substitue la valeur de $u_r$ obtenue en (2.26)

$$k_p = k_q \frac{A_p}{A_q} + P_q \frac{A_p}{A_q} - P_q \tag{2.27}$$

En prenant la somme de ces équations pour $p \in Q_r - \{q\}$

$$d^r - k_q = \sum_{p \in Q_r - \{q\}} k_p = k_q \left(\frac{1-A_q}{A_q}\right) + P_q \left(\frac{1-A_q}{A_q}\right) - \sum_{p \in Q_r - \{q\}} P_p$$

D'où

$$k_q = A_q\, d^r - (1-A_q)\, P_q + A_q \sum_{p \in Q_r - \{q\}} P_p \tag{2.28}$$

La relation (2.28) peut être interprétée de la façon suivante. $A_q\ d^r$ est la part de la demande originellement attirée par la route q. Le nombre de passagers excédentaires retranchés est $(1-A_q)P_q$ et non $P_q$. La différence $A_q\ P_q$ est le biais résultant de l'hypothèse oubliante; elle correspond aux passagers excédentaires qui réessaient cette même route. Le dernier terme correspond à la partie des passagers refusés sur les autres routes qui se retrouvent sur la route q.

## 2.7 Modèle P4: Répartition des passagers en considérant l'"optimisation égoïste" et une demande élastique.

Ce modèle traite simultanément le problème du choix des routes par les passagers et l'équilibre offre/demande. Ces deux problèmes sont en relation car le niveau de service associé à l'offre influence la demande. En effet un bon service attire une plus grande partie des passagers potentiels. Le niveau de service offert à un marché dépend des routes que peuvent utiliser ces passagers. Des routes très attrayantes offrent peu de service à un marché si elles sont déjà saturées par des passagers d'autres marchés.

Précisons que le niveau de service dépend de deux facteurs: l'attraction et la disponibilité des routes offertes. L'attraction des routes offertes est connue à priori et est indépendante de la répartition des passagers. Ainsi, l'effet de ce facteur sur la demande peut être traité de façon externe au modèle. On peut définir une fonction de l'attraction des routes offertes donnant la partie des clients potentiels qui tenteront de voyager. Ce facteur devient important quand on change les routes offertes, il sera traité au chapitre 5. Par contre, ceux qui tentent de voyager ne trouvent pas tous une place acceptable; certains sont découragés car les routes qui les ont attirés ne sont pas disponibles. La disponibilité des routes n'est pas connue

a priori, elle dépend de la répartition des passagers. Ainsi l'effet de la disponibilité sur la demande doit être traité simultanément avec la répartition des passagers.

En résumé la répartition des passagers dépend de la demande qui dépend de la répartition des passagers par l'effet du facteur disponibilité sur le niveau de service.

## 2.7.1 Mesure de la disponibilité des routes

Il faut associer à la répartition des passagers une variable indiquant l'état de disponibilité des routes. Revenons à (2.19) l'équation fondamentale du modèle P3.

$$a_q = k_q + P_q(k) = u_r A_q \text{ pour tout } q \in Q_r$$

$a_q$ indiquant le nombre total de passagers essayant d'utiliser la route q, $k_q$ ceux qui y trouvent place et $P_q(k)$ ceux refusés faute de place. En prenant la somme sur toutes les routes $q \in Q_r$ nous obtenons

$$\sum_{q \in Q_r} a_q = \sum_{q \in Q_r} k_q + \sum_{q \in Q_r} P_q(k) = u_r \sum_{q \in Q_r} A_q = u_r \qquad (2.29)$$

Cette équation indique la signification de $u_r$ qui correspond au nombre total d'essais en vue d'utiliser une route que les passagers du marché r ont effectués. Ainsi, pour une demande constante $\left(\sum_{q \in Q_r} k_q = d^r\right)$, $u_r$ varie comme le nombre d'essais infructueux $\left(\sum_{q \in Q_r} P_q(k)\right)$ en vue d'utiliser une des routes. Il reflète donc l'inconvénient causé aux passagers par le manque de disponibilité. Il faut donc s'attendre à trouver une relation entre $u_r$ et le nombre de passagers qui voyagent.

2.7.2 Relation entre le nombre de passagers et la disponibilité des routes

Nous allons établir une relation entre les quantités suivantes:

$v^r$ = le nombre variable de passagers du marché r qui voyagent

$d^r$ = le nombre de clients qui tentent de voyager. Ce nombre dépend du marché et de l'attraction des routes mais est constant pendant l'assignation des passagers. ($d^r > 0$ sinon le marché n'existe pas)

$\lambda_r$ = la proportion des clients du marché r qui choisissent une autre route s'ils sont refusés, faute de place. Les autres choisissent de ne pas voyager.

$u_r$ = la variable duale du problème d'assignation des passagers dont on a discuté à la section 2.7.1.

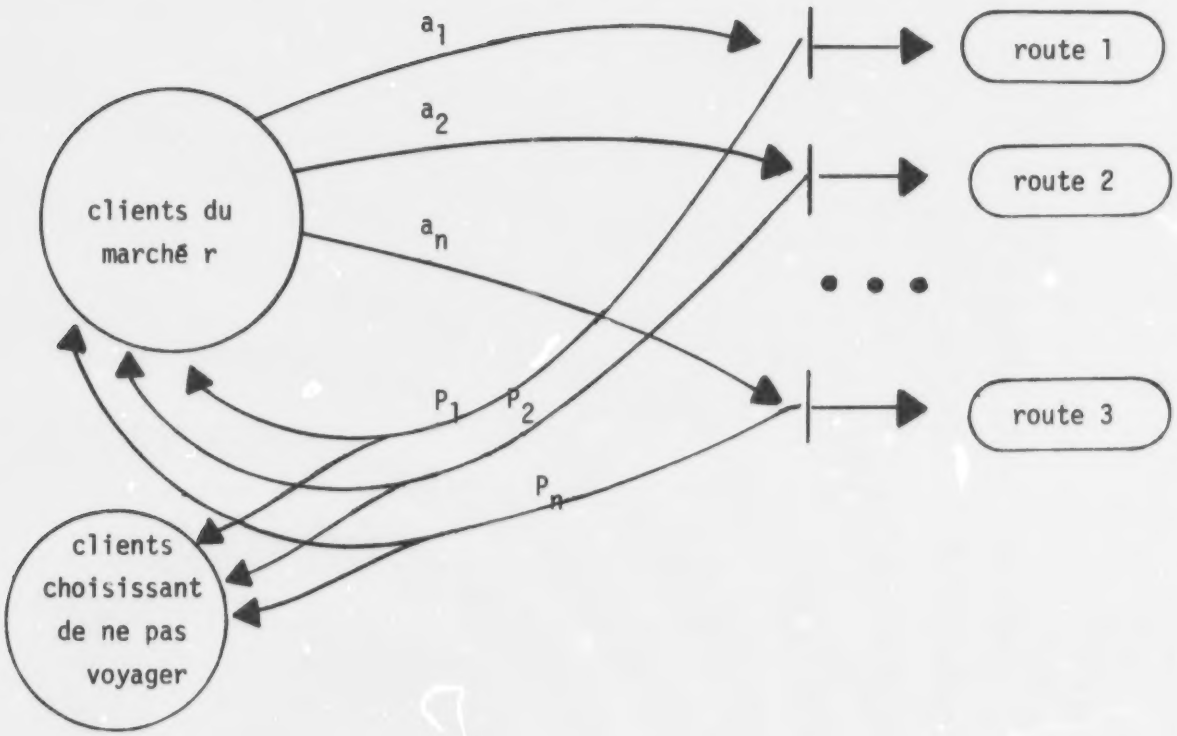
Pour définir cette relation, revenons au cheminement des passagers suivant l'hypothèse oubliante et modifions-la. La figure 12 présente le cheminement modifié en considérant qu'une partie des passagers refusés faute de place choisissent de ne pas voyager.

Le total ( $\sum_{q \in Q_r} a_q$ ) des passagers qui essaient de voyager sur les n routes étant constitué des $d^r$ clients et de la partie des passagers refusés qui essaient une nouvelle route nous obtenons la relation suivante:

$$\sum_{q \in Q_r} a_q = d_r + \lambda_r \sum_{q \in Q_r} P_q(k) \qquad (2.30)$$

FIGURE 12

CHEMINEMENT DES PASSAGERS QUAND UNE PARTIE DES PASSAGERS REFUSÉS FAUTE DE PLACE, CHOISISSENT DE NE PAS VOYAGER




**En y substituant (2.29) et (2.19) cette équation devient**

$$u_r = d^r + \lambda_r \left(u_r - \sum_{q \in Q_r} k_q\right) \qquad (2.31)$$

Comme $\sum_{q \in Q_r} k_q = v^r$ nous obtenons la relation cherchée entre $v^r$ et $u_r$

$$v^r(u_r) = d^r - \left(\frac{1-\lambda_r}{\lambda_r}\right)(u_r - d^r) \qquad (2.32)$$

On peut remarquer que $v^r$ atteint sa valeur maximum $d^r$ lorsque $u_r = d^r$. La relation (2.29) permet de voir que dans ce cas aucun passager n'est refusé faute de place ( $\sum_{q \in Q_r} P_q(k) = 0$). Quand $u_r$ augmente, le nombre de passagers refusés faute de place augmente et il faut donc que $v^r$ diminue. La fonction de demande $v^r$ peut aussi être développée dans le cas plus général où les $\lambda_r$ sont remplacés par des $\lambda_{r,q}$ dépendant de la route. Dans ce cas $v^r$ est fonction de $u_r$ et des $k_q$; le traitement devient plus complexe mais quand même possible.

### 2.7.3 Equations du modèle

Afin d'intégrer l'équation de demande variable (2.32) au modèle, nous la transformons de façon à exprimer $u_r$ en fonction de $v^r$. La relation suivante s'obtient facilement car (2.32) est linéaire.

$$u_r = \phi(v^r) = d^r + \left( \frac{\lambda_r}{1-\lambda_r} \right)(d^r - v^r) \qquad (2.33)$$

Remarquons que le cas $v^r = 0$ ne peut se produire. En effet si $v^r = 0$ aucun passager voyage, alors $k_q = 0$ pour tout $q \in Q_r$. De là, le nombre de passagers excédentaires $P_q(k)$ est aussi nul suivant sa définition donnée à l'équation (2.11). Donc $u_r = 0$ car par l'équation (2.19) $k_q + P_q(k) = u_r A_q$ et $A_q > 0$. Ceci est contradictoire car l'équation (2.32) permet de calculer que $v^r > 0$ quand $u_r = 0$

On peut donc remplacer (2.33) par

$$\phi(v^r) \quad \begin{cases} = u_r & \text{si } v^r > 0 \\ \leq u_r & \text{si } v^r = 0 \end{cases} \qquad (2.34)$$

Ainsi, en ajoutant la relation de demande variable (2.34) aux équations d'équilibre (2.21) du modèle précédent, nous obtenons les conditions d'équilibre suivantes:

Pour tout r il existe $u_r$ tel que

$$\frac{k_q + P_q(k)}{A_q} \quad \begin{cases} = u_r & \text{si } k_q > 0 \\ \geq u_r & \text{si } k_q = 0 \end{cases} \qquad (\forall q \in Q_r)$$

$$\phi\,(v^r) \quad \begin{cases} = u_r & \text{si } v^r > 0 \\ \leq u_r & \text{si } v^r = 0 \end{cases}$$

$$\sum_{q \in Q_r} k_q - v^r = 0$$

$$k_q \geq 0 \qquad (\forall q \in Q_r)$$

$$v^r \geq 0 \qquad (2.35)$$

Ce système présente la même structure que (2.21) si on exclut le signe négatif de la variable $v^r$. Ce problème se résoud de façon semblable à celui où la demande est fixe. L'existence de solution et l'adaptation des algorithmes de résolutions sont discutés au chapitre 3.

### 2.7.4 Propriétés du modèle

Un modèle reproduisant l'équilibre offre/demande est très utile pour le planificateur. Un tel modèle permet d'estimer de façon réaliste les revenus espérés d'un projet d'horaire. L'effet d'un mauvais service pour certains marchés est quantifié par le modèle en terme de perte de revenus. L'analyse d'un horaire avec les modèles précédents fournissait des indices de satisfactions, des taux de passagers excé-

dentaires mais ne donnait pas la mesure de qualité très importante qu'est le revenu espéré.

La fonction de demande utilisée par ce modèle est facile à calibrer, elle comprend qu'un seul paramètre par marché et ce paramètre $\lambda_r$ est directement observable. Les $\lambda_r$ peuvent être obtenus du système de réservation en enregistrant les clients qui n'achètent pas de billet si le vol désiré n'a plus de places disponibles.

CHAPITRE 3

# ALGORITHME DE RESOLUTION DU PROBLEME DES PASSAGERS

Les quatre modèles proposés donnent lieu à des structures de problèmes semblables. Dans chaque cas les contraintes formant des blocs diagonaux disjoints, chaque bloc reliant la demande et les chemins d'un marché. Dans chaque cas également les conditions d'optimalité lient tous les marchés. Pour les modèles P1 et P2,l'optimalité correspond à minimiser un objectif global ne se séparant pas par marché. Rappelons la forme de ces problèmes

$$\text{Min } f(h)$$

$$\sum_{q \in Q_r} h_q = d^r \qquad (\forall r)$$

$$h_q \geq 0 \qquad (\forall q \in Q_r) \qquad (\forall r) \qquad (3.1)$$

Pour les problèmes P3 et P4,l'optimalité est définie par des équations d'équilibre ne se séparant pas par marché. Ces problèmes ont la forme suivante.

Pour tout r il existe un $u_r$ tel que:

$$C_q(k) \quad \left\{ \begin{array}{ll} = u_r & \text{si} \quad k_q > 0 \\ \geq u_r & \text{si} \quad k_q = 0 \end{array} \right\} \quad (\forall q \in Q_r)$$

$$\sum_{q \in Q_r} k_q = d^r \qquad (\forall r)$$

$$k_q \geq 0 \qquad (\forall q \in Q_r) \qquad (\forall r) \qquad (3.2)$$

Rappelons que $C_q(k) = \left(k_q + P_q(k)\right) / A_q$. Remarquons que les équations (2.35) qui définissent le modèle P4 présentent aussi cette structure en considérant une route suplémentaire à laquelle est associée la variable $v^r$. Le coefficient négatif de $v^r$, dans la contrainte de partage entre les routes $(\Sigma q \epsilon Q_r \ k_q - v^r = 0)$, ne complique pas la résolution comme nous le verrons par la suite.

La ressemblance entre la structure du problème 3.1 et celle de 3.2 devient plus évidente si on remplace 3.1 par les conditions d'optimalités de Kuhn-Tucker.

Pour tout r il existe $u_r$ tel que:

$$\frac{\partial f(h)}{\partial h_q} \begin{cases} = u_r & \text{si} \quad h_q > 0 \\ \geq u_r & \text{si} \quad h_q = 0 \end{cases} \quad (\forall q \epsilon Q_r)$$

$$\sum_{q \epsilon Q_r} h_q = d^r \qquad (\forall r)$$

$$h_q \geq 0 \qquad (\forall q \epsilon Q_r) \qquad (\forall r) \qquad (3.3)$$

Il convient de rappeler, qu'il n'existe pas en général de fonction f permettant d'écrire le système 3.2 sous la forme 3.1 . Le système 3.1 est donc un cas particulier du système 3.2 .

Plutôt que de nous attaquer directement à la résolution du cas général, nous allons d'abord discuter d'une méthode de résolution du cas particulier 3.1 . Par la suite, nous adapterons l'idée de cette méthode au cas général. Il y a avantage à traiter P1 et P2 comme des problèmes d'optimisation car les conditions d'existence, d'unicité de solution et de convergence sont ainsi plus faciles à vérifier. Les modèles P3 et P4

doivent être traités comme des systèmes d'équations non-linéaires et des conditions plus fortes doivent être vérifiées pour assurer l'existence et l'unicité de solution et la convergence de l'algorithme.

## 3.1 Décomposition du problème

Le problème 3.1 est résolu par décomposition en sous-problèmes correspondant à chaque marché. L'ensemble des marchés est parcouru de façon cyclique, et, un à un, les sous-problèmes sont réoptimisés en maintenant fixes les variables des autres marchés. Introduisons d'abord la notation nécessaire pour préciser la procédure. Une affectation h des passagers est dite réalisable si elle satisfait les contraintes du système 3.1 . Soit $h_{Q_r} = (h_{q_1}, h_{q_2}, \ldots, h_{q_n})$ où $\{ q_1, \ldots, q_n\}$ est l'ensemble der routes desservant le marché r.

Soit $S_r$ la fonction qui correspond à réoptimiser le sous-problème du marché r tout en maintenant fixes les variables des autres marchés

$$S_r: (h_{Q_1}, \ldots, h_{Q_m}) \rightarrow (h_{Q_1}, \ldots, h_{Q_{r-1}}, h^*_{Q_r}, h_{Q_{r+1}}, \ldots h_{Q_m}) \qquad (3.4)$$

où m est le nombre de marché et $h^*_{Q_r}$ la solution optimale du sous-problème r les autres variables étant fixées.

$$\text{Soit } S = S_m o \; S_{m-1} o \; \ldots o \; S_1 \; . \qquad (3.5)$$

Cette fonction correspond à réoptimiser successivement chacun des sous-problèmes.

L'algorithme global de résolution consiste à débuter avec $h^0$ une solution réalisable et à itérer la fonction S en définissant

$$h^{k+1} = S(h^k) \; . \qquad (3.6)$$

### 3.1.1 Convergence de la méthode de décomposition pour les modèles P1 et P2

L'existence de solution ne pose pas de problèmes. Le domaine réalisable est compact, étant fermé et borné. Comme il est non vide, la fonction f y atteint son minimum.

Discutons maintenant l'unicité de la solution. Il est bien connu que quand la fonction f est strictement convexe, le point minimum est unique. Nous avons constaté à la section 2.4.3 que la composante des objectifs des modèles P1 et P2 qui est associée à la limite de capacité des avions, peut être non-convexe dans certains cas. Cependant, nous avons remarqué que les problèmes réels sont probablement convexes et même strictement convexes.

La convergence de la procédure de décomposition pour le problème d'optimisation 3.1 est montrée en se référant aux résultats de Zangwill (69). Précisons d'abord les termes: Soit S une application de V dans lui-même associant à tout point $v \in V$ un ensemble $S(v) \subseteq V$. Cette application est fermée, si pour tout $v \in V$

a) $v^k \to v$ , $v^k \in V$

b) $\omega^k \to \omega$ , $\omega^k \in S(v^k)$

impliquent que

c) $\omega \in S(v)$

L'application S génère à partir d'un $h^0$ une suite de points $(h^k)$ diminuant la valeur de la fonction économique f à chaque itération, si elle satisfait les conditions suivantes:

a) $h^{k+1} \in S(h^k)$

b) si $h^k$ n'est pas solution optimale alors pour tout $Y \in S(h^k)$ on a $f(Y) < f(h^k)$

c) si $h^k$ est une solution optimale ou bien $S(h^k) = \phi$ ou bien pour tout $Y \in S(h^k)$ on a $f(Y) \geq f(h^k)$

Zangwill a montré que si S est fermé et génère à partir d'un $h^o$ une suite diminuant la valeur de la fonction économique à chaque itération, alors tout point d'accumulation de cette suite est une solution optimale.

Ce théorème s'applique à notre procédure de décomposition en prenant V égale à l'ensemble des affectations réalisables et en prenant l'application S tel que définie à l'équation 3.6, ainsi $S(h^k) = \{h^{k+1}\}$. Pour montrer que S est fermé, il suffit de vérifier que chacun des $S_r$ est fermé car Zangwill a montré que la composition d'application fermée est fermée.

Lemme 1: $S_r$ est fermé

Preuve: afin de réduire la notation prenons

$$(h_{Q_r}, h_{\bar{Q}_r}) = (h_{Q_1}, h_{Q_2}, \ldots, h_{Q_r}, \ldots, h_{Q_n})$$

où $\bar{Q}_r$ est l'ensemble de toutes les routes hors de $Q_r$ .

Il faut montrer que si

a) $(h^k_{Q_r}, h^k_{\bar{Q}_r}) \rightarrow (h_{Q_r}, h_{\bar{Q}_r})$

b) $(\omega^k_{Q_r}, \omega^k_{\bar{Q}_r}) = (h^{*k}_{Q_r}, h^k_{\bar{Q}_r}) \rightarrow (\omega_{Q_r}, \omega_{\bar{Q}_r})$

Alors,

$$c) \quad (\omega_{Q_r}, \omega_{\bar{Q}_r}) = (h^*_{Q_r}, h_{\bar{Q}_r}) \ .$$

D'abord puisque $h^k_{\bar{Q}_r} \to h_{\bar{Q}_r}$, alors $\omega_{\bar{Q}_r} = h_{\bar{Q}_r}$ .

Rappelons que par définition de $S_r$ la solution optimale du sous-problème r est $h^{*k}_{Q_r}$ lorsqu'on fixe les autres variables égales à $h^k_{\bar{Q}_r}$, d'où

$$f(h^{*k}_{Q_r}, h^k_{\bar{Q}_r}) \leq f(x_{Q_r}, h^k_{\bar{Q}_r}) \ (\forall x_{Q_r}) \ . \tag{3.7}$$

De même $h^*_{Q_r}$ est la solution optimale du sous-problème r lorsqu'on fixe les autres variables égales à $h_{\bar{Q}_r}$. Donc

$$f(h^*_{Q_r}, h_{\bar{Q}_r}) \leq f(x_{Q_r}, h_{\bar{Q}_r}) \ \ (\forall x_{Q_r}) \ . \tag{3.8}$$

Démontrons maintenant que $\omega_{Q_r} = h^*_{Q_r}$ en établissant la relation 3.8 avec $\omega_{Q_r}$ à la place de $h^*_{Q_r}$

$$f(\omega_{Q_r}, h_{\bar{Q}_r}) = f(\omega_{Q_r}, h_{\bar{Q}_r}) - f(h^{*k}_{Q_r}, h^k_{\bar{Q}_r}) + f(h^{*k}_{Q_r}, h^k_{\bar{Q}_r}) \ .$$

Puisque f est continu et que $(h^{*k}_{Q_r}, h^k_{\bar{Q}_r}) \to (\omega_{Q_r}, h_{\bar{Q}_r})$ .

Alors pour k suffisamment grand

$$f(\omega_{Q_r}, h_{\bar{Q}_r}) \leq \epsilon + f(h^{*k}_{Q_r}, h^k_{\bar{Q}_r}) \ .$$

En utilisant la relation 3.7 l'inéquation devient

$$f(\omega_{Q_r}, h_{\bar{Q}_r}) \leq \epsilon + f(x_{Q_r}, h^k_{\bar{Q}_r}) \quad (\forall x_{Q_r}),$$

ou encore

$$f(\omega_{Q_r}, h_{\bar{Q}_r}) \leq \epsilon + f(x_{Q_r}, h^k_{\bar{Q}_r}) - f(x_{Q_r}, h_{\bar{Q}_r}) + f(x_{Q_r}, h_{\bar{Q}_r}) \quad (\forall x_{Q_r}) \ .$$

Comme f est uniformément continue sur l'ensemble compact des solutions réalisables et que $h^k_{\bar{Q}_r} \to h_{\bar{Q}_r}$,

alors pour k suffisamment grand

$$f(\omega_{Q_r}, h_{\bar{Q}_r}) \leq \epsilon + \epsilon + f(x_{Q_r}, h_{\bar{Q}_r}) \quad (\forall x_{Q_r})$$

Comme cette relation est satisfaite pour tout $\epsilon$ il faut que

$$f(\omega_{Q_r}, h_{\bar{Q}_r}) \leq f(x_{Q_r}, h_{\bar{Q}_r}) \quad (\forall x_{Q_r}) \ . \tag{3.9}$$

Donc $(\omega_{Q_r}, \omega_{\bar{Q}_r}) = (h^*_{Q_r}, h_{\bar{Q}_r})$ et ainsi $S_r$ est fermé.

Montrons maintenant que l'algorithme vérifie la seconde hypothèse du théorème.

Lemme 2. S diminue la valeur de la fonction économique à chaque itération.

Preuve: Si $h^k$ n'est pas une solution optimale, alors il existe un premier sous-problème r pour lequel elle n'est pas optimale. Ceci découle du fait que les conditions de Kuhn-Tucker du problème global sont l'union des conditions d'optimalité de Kuhn-Tucker des sous-problèmes.

L'application $S_1, S_2, \ldots, S_{r-1}$ ne modifie donc pas la solution. Ainsi

$$h^k = S_1(h^k) = \ldots = S_{r-1} o \ldots o S_2 \, o S_1(h^k) \quad .$$

L'application de $S_r$ génère une nouvelle solution ayant une valeur de f inférieure. Ainsi

$$f\left(S_r o \, S_{r-1} o \ldots o S_1(h^k)\right) < f\left(S_{r-1} o \ldots o S_1(h^k)\right) = f(h^k) \quad .$$

L'application de $S_{r+1}, \ldots, S_m$ génère des solutions n'augmentant pas la valeur de f. Donc

$$f(h^{k+1}) = f\left(S_m o \ldots o S_1(h^k)\right) \leq f\left(S_r o \ldots o S_1(h^k)\right) < f(h^k) \quad .$$

D'autre part si $h^k$ est solution optimale elle n'est pas modifiée lors de l'application de chacun des $S_1, S_2, \ldots, S_m$. Donc

$$h^{k+1} = S(h^k) = h^k \quad \text{et} \quad f(h^{k+1}) = f(h^k) \quad .$$

En pratique les sous-problèmes sont résolus seulement à $\epsilon$ près car leur résolution exacte demande un nombre infini d'itérations. Etudions la convergence de la méthode de décomposition dans ce contexte.

Dans ce cas l'image de la fonction $S_r$ est l'ensemble suivant.

$$S_r(h_{Q_r}, h_{\bar{Q}_r}) = \left\{ (x_{Q_r}, h_{\bar{Q}_r}) \;\middle|\; |f(x_{Q_r}, h_{\bar{Q}_r}) - f(h^*_{Q_r}, h_{\bar{Q}_r})| \leq \epsilon \right\} \qquad (3.10)$$

où $h^*_{Q_r}$ est la solution optimale comme en 3.4.

Le théorème de Zangwill montre la convergence vers une solution qui satisfait simultanément tous les sous-problèmes à $\epsilon$ près. Pour appliquer ce théorème, il faut établir les lemmes montrant que S est fermé et génère une suite diminuant la valeur de la fonction économique à chaque itération.

Lemme 3. $S_r$ définie par la relation 3.10 est fermé.

Preuve: Il faut montrer que si

a) $(h^k_{Q_r}, h^k_{\overline{Q}_r}) \to (h_{Q_r}, h_{\overline{Q}_r})$

b) $(\omega^k_{Q_r}, \omega^k_{\overline{Q}_r}) \to (\omega_{Q_r}, \omega_{\overline{Q}_r})$

où $|f(\omega^k_{Q_r}, \omega^k_{\overline{Q}_r}) - f(h^{*k}_{Q_r}, h^k_{\overline{Q}_r})| \leq \epsilon$ . (3.11)

Alors

c) $|f(\omega_{Q_r}, \omega_{\overline{Q}_r}) - f(h^*_{Q_r}, h_{\overline{Q}_r})| \leq \epsilon$ . (3.12)

Remarquons que

$$|f(\omega_{Q_r}, \omega_{\overline{Q}_r}) - f(h^*_{Q_r}, h_{\overline{Q}_r})|$$

$$\leq |f(\omega_{Q_r}, \omega_{\overline{Q}_r}) - f(\omega^k_{Q_r}, \omega^k_{\overline{Q}_r})| + |f(\omega^k_{Q_r}, \omega^k_{\overline{Q}_r}) - f(h^{*k}_{Q_r}, h^k_{\overline{Q}_r})| + |f(h^{*k}_{Q_r}, h^k_{\overline{Q}_r}) - f(h^*_{Q_r}, h_{\overline{Q}_r})|$$

$$\leq \quad \rho \quad + \quad \epsilon \quad + \quad \rho \ .$$

En effet en prenant k suffisamment grand, le premier terme est plus petit que $\rho$ étant donné que f est continue et que

$(\omega^k_{Q_r}, \omega^k_{\overline{Q}_r}) \to (\omega_{Q_r}, \omega_{\overline{Q}_r})$ .

Le second terme est plus petit que $\epsilon$ par 3.11 .

Le troisième terme est aussi plus petit que $\rho$ en prenant k suffisamment grand car f est uniformément continue et $(\forall x_{Q_r})$, $(x_{Q_r}, h^k_{\overline{Q}_r}) \to (x_{Q_r}, h_{\overline{Q}_r})$

d'où

$$|f(h^{*k}_{Q_r}, h^{k}_{\overline{Q}_r}) - f(h^{*}_{Q_r}, h_{\overline{Q}_r})| \leq \underset{x_{Q_r}}{\text{Max}} |f(x_{Q_r}, h^{k}_{\overline{Q}_r}) - f(x_{Q_r}, h_{\overline{Q}_r})| \leq \rho \quad .$$

Donc nous avons 3.12 car pour tout $\rho > 0$

$$|f(\omega_{Q_r}, \omega_{\overline{Q}_r}) - f(h^{*}_{Q_r}, h_{\overline{Q}_r})| \leq 2\rho + \epsilon \quad .$$

Lemme 4 S définie par la relation 3.10 diminue la fonction économique à chaque itération.

Preuve: La preuve est similaire à celle du lemme 2. Il suffit de montrer si $h^k$ n'est pas optimal à $\epsilon$ près pour un sous-problème r, la fonction f est diminuée lors de l'application de $S_r$ .

Si le sous-problème r n'est pas optimal à $\epsilon$ près

$$|f(h^{k}_{Q_r}, h^{k}_{\overline{Q}_r}) - f(h^{*k}_{Q_r}, h^{k}_{\overline{Q}_r})| \geq \epsilon$$

$$\text{d'où} \quad f(h^{k}_{Q_r}, h^{k}_{\overline{Q}_r}) \geq f(h^{*k}_{Q_r}, h^{k}_{\overline{Q}_r}) + \epsilon \quad . \tag{3.13}$$

Par contre $S_r(h^{k}_{Q_r}, h^{k}_{\overline{Q}_r})$ est optimal à $\epsilon$ près

$$|f(S_r(h^{k}_{Q_r}, h^{k}_{\overline{Q}_r})) - f(h^{*k}_{Q_r}, h^{k}_{\overline{Q}_r})| < \epsilon$$

$$\text{d'où} \quad f(S_r(h^{k}_{Q_r}, h^{k}_{\overline{Q}_r})) < f(h^{*k}_{Q_r}, h^{k}_{\overline{Q}_r}) + \epsilon \quad . \tag{3.14}$$

En combinant 3.13 et 3.14 on obtient le résultat désiré

$$f(S_r(h^{k}_{Q_r}, h^{k}_{\overline{Q}_r})) < f(h^{k}_{Q_r}, h^{k}_{\overline{Q}_r}) \quad .$$

3.1.2 Existence de solutions du problème P3.

Nous présentons ici une preuve d'existence exigeant seulement des conditions relativement faibles et facilement vérifiables. Le théorème de convergence présenté à la section 3.1.4 démontrera également l'existence de la solution, mais les conditions de ce théorème sont difficilement vérifiables. La discussion permettra de voir que ces conditions sont généralement réalisées même si nous ne pouvons en faire la preuve dans le cas général.

Cette première preuve d'existence de solution utilise le résultat obtenu par Karamardian (69) à partir du théorème de point fixe de Kakutani. Soit V un sous-ensemble de $E^n$, compact, convexe non vide, et soit C: $V \to E^n$ une fonction continue. Alors il existe $\bar{k} \in V$ tel que

$$(k - \bar{k})\, C(\bar{k}) \geq 0 \quad \text{pour tout } k \in V \ . \tag{3.15}$$

Ce théorème s'applique en prenant V l'ensemble des solutions réalisables de 3.2 (ensemble compact et convexe) et C le vecteur de fonction continue tel que défini dans le problème 3.2. Rappelons ici que C a été prolongé linéairement à l'annexe 3 de façon à être défini et continue sur tout l'ensemble V.

Montrons maintenant que le $\bar{k}$ dont l'existence est assuré par le théorème de Karamardian satisfait les conditions d'optimalité du problème 3.2 .

D'abord comme $\bar{k}$ est réalisable les conditions suivantes sont satisfaites.

$$\sum_{q \in Q_r} \bar{k}_q = d^r \qquad (\forall r)$$

$$\bar{k}_q \geq 0 \quad (\forall q \in Q_r) \quad (\forall r) \ .$$

Prenons $u_r = \min_{q \in Q_r} C_q(\bar{k})$ .

Cette définition assure que $C_q(\bar{k}) \geq u_r$ en particulier quand $\bar{k}_q = 0$

Pour r quelconque prenons $p \in Q_r$ tel que $C_p(\bar{k})$ soit minimum

i.e. $(C_p(\bar{k}) = u_r)$ .

Prenons $k \in V$ défini de la façon suivante:

$k_q = \bar{k}_q$ si $q \in Q_s$ et $s \neq r$

$k_q = 0$ si $q \in Q_r$ et $q \neq p$

$k_p = d^r$ .

En utilisant l'équation 3.15, on déduit que

$$d^r u_r - \sum_{q \in Q_r} \bar{k}_q C_q(\bar{k}) \geq 0 .$$

D'où $\sum_{q \in Q_r} \bar{k}_q \left(C_q(\bar{k}) - u_r\right) \leq 0$ car $\sum_{q \in Q_r} \bar{k}_q = d^r$ .

Donc $\bar{k}_q \left(C_q(\bar{k}) - u_r\right) = 0 \quad (\forall q \in Q_r)$ .

car une somme de termes non négatifs $\left(\bar{k}_q \geq 0 \text{ et } C_q(\bar{k}) - u_r \geq 0\right)$ est non-positive seulement si tous les termes sont nuls.

De là si $\bar{k}_q > 0$ il faut que $C_q(\bar{k}) = u_r$ . Ceci démontre que $\bar{k}$ satisfait les conditions d'optimalité du problème 3.2 .

### 3.1.3 Existence de solution pour le problème P4

Cette preuve d'existence fait également appel au résultat de Karamardian déjà introduit. Ici V est l'ensemble des solutions réalisables du problème 2.35. Comme le modèle traite une demande variable, ces solutions sont formées des variables $v^r$ en plus des variables $k_q$.

V est compact car il est fermé et borné. En effet, il est borné

$$0 \leq k_q \leq v^r \leq d^r$$

s'obtient des équations du problème (2.35) et de l'évaluation de $v^r$ par l'équation (2.32). La convexité de V découle directement des équations linéaires qui définissent les solutions réalisables.

$$\left( \sum_{q \in Q_r} k_q - v^r = 0,\ v^r \geq 0,\ (h_q \geq 0,\ \forall q \in Q_r) \right)\ (\forall r)\ .$$

La fonction C comprend les composantes $C_q$ communes au problème 3.2 et comprend en plus la composante $-\phi(v^r)$ définie par la relation 2.33.
C est continue car toutes ses composantes le sont.

Dans ces conditions le théorème de Karamardian assure l'existence de $(\bar{k}, \bar{v}) \in V$ tel que

$$((k, v) - (\bar{k}, \bar{v}))\ C(\bar{k}\ \bar{v}) \geq 0 \quad \forall (k, v) \in V \qquad (3.16)$$

Montrons maintenant que $(\bar{k}, v)$ satisfait les conditions (2.35).

Les conditions linéaires sont satisfaites $(\bar{k}, v)$ étant une solution réalisable

Soit $u_r = \min_{q \in Q_r} C_q(\bar{k})$

Cette définition assure que $C_q(\bar{k}) \geq u_r$ en particulier quand $\bar{k}_q = 0$

Pour r quelconque prenons $p \in Q_r$ tel que $C_p(\bar{k})$ soit minimum,

i.e. $(C_p(\bar{k}) = u_r)$ .

Prenons la variable $x \in [0, d^r)$ et définissons $(k, v) \in V$ comme suit

$k_q = \bar{k}_q$ si $q \in Q_s$ et $s \neq r$

$k_q = 0$ si $q \in Q_r$ et $q \neq p$

$k_p = v^r = x$ .

En utilisant l'équation (3.16) on déduit que pour tout $x \in [0, d^r]$

$$x\, u_r - x\, \phi(\bar{v}^r) - \sum_{q \in Q_r} \bar{k}_q\, C_q(\bar{k}) + \bar{v}^r\, \phi(\bar{v}^r) \geq 0 . \qquad (3.17)$$

En particulier pour $x = \bar{v}^r$, (3.17) devient

$$\bar{v}^r\, u_r - \sum_{q \in Q_r} \bar{k}_q\, C_q(\bar{k}) \geq 0 .$$

D'où $\sum_{q \in Q_r} \bar{k}_q \left(C_q(\bar{k}) - u_r\right) \leq 0$ car $\bar{v}^r = \sum_{q \in Q_r} \bar{k}_q$ .

Donc $\bar{k}_q \left(C_q(\bar{k}) - u_r\right) = 0 \quad (\forall q \in Q_r)$

car une somme de termes non négatifs $\left(\bar{k}_q \geq 0 \text{ et } C_q(\bar{k}) - u_r \geq 0\right)$ est non positive seulement si tous ses termes sont nuls.

Ceci permet d'obtenir une des conditions d'optimalité

Si $\bar{k}_q > 0$ alors $C_q(\bar{k}) = u_r$

Transformons (3.17) avec x variable en utilisant ce dernier résultat

$$x\, u_r - x\, \phi(\bar{v}^r) - \sum_{q \in Q_r} \bar{k}_q\, u_r + \bar{v}^r\, \phi(\bar{v}^r) \geq 0$$

D'où $(x-\bar{v}^r)u_r - (x-\bar{v}^r)\,\phi(\bar{v}^r) \geq 0$ car $\sum_{q\in Q_r} \bar{k}_q = \bar{v}^r$ .

Donc $(x-\bar{v}^r)\,(u_r - \phi(\bar{v}^r)) \geq 0$ . (3.18)

En prenant $x = d^r$ dans la relation (3.18)

$(d^r - \bar{v}^r)\,(u_r - \phi(\bar{v}^r)) \geq 0$ .

Dans le cas où $d^r > \bar{v}^r$ il est immédiat que $\phi(\bar{v}^r) \leq u_r$ .

Dans le cas où $d^r = \bar{v}^r$, $\phi(\bar{v}^r) = d^r$ par (2.33) et

$d^r = \bar{v}^r = \sum_{q\in Q_r} \bar{k}_q \leq \sum_{q\in Q_r} \bar{k}_q + \sum_{q\in Q_r} P_q(\bar{k}) = u_r$ par (2.29) .

Donc dans ce cas aussi $\phi(\bar{v}^r) \leq u_r$ .

Ceci assure la condition d'optimalité

Si $\bar{v}^r = 0$ alors $\phi(\bar{v}^r) \leq u^r$ .

Montrons maintenant la condition d'optimalité suivante:

Si $\bar{v}^r > 0$ alors $\phi(\bar{v}^r) = u^r$ .

Si $\bar{v}^r > 0$ en prenant $x = 0$ la relation (3.18) devient

$-\bar{v}^r\,(u_r - \phi(\bar{v}^r)) \geq 0$ .

D'où $\phi(\bar{v}^r) \geq u_r$ .

Donc $\phi(\bar{v}^r) = u_r$ .

Ceci complète la vérification des conditions (2.35).

Donc $(\bar{k}, \bar{v})$ est une solution optimale du problème P4.

### 3.1.4 Convergence de la méthode de décomposition pour les modèles P3.

Dans ce cas, la théorie de Zangwill ne peut être utilisée pour montrer la convergence de la méthode de résolution car il n'existe pas d'objectif qui décroît d'une itération à l'autre. Par contre, la même méthode de résolution peut être utilisée et être considérée dans ce contexte plus général comme une extension de la méthode de Gauss-Siedel pour la résolution des systèmes d'équations non-linéaires.

Revenons aux équations originales (2.20). En considérant $A_q > 0$ et en définissant encore $C_q(k) = \left(k_q + P_q(k)\right)/A_q$ ce système d'équations devient:

pour tout r il existe $u_r$ tel que:

$$C_q(k) = u_r \qquad (\forall q \in Q_r)$$

$$\sum_{q \in Q_r} k_q = d^r \tag{3.19}$$

Ce système comprend m groupes de $n_r$ équations où m est le nombre de marchés et $n_r = |Q_r| + 1$ le nombre d'équations pour chaque marché. La méthode de résolution proposée consiste à réajuster un à un chaque groupe de $n_r$ variables en résolvant le système d'équations du marché correspondant. Cette procédure de résolution par bloc est une extension de la méthode de Gauss-Siedel pour les systèmes non-linéaires qui réajuste une à une les variables en utilisant une équation à la fois.

La convergence de la méthode de Gauss-Siedel par bloc a été étudiée par Varga (62) dans le cas linéaire. Pour le problème AX = b où

la matrice A présente la structure de bloc suivante

$$A = \begin{pmatrix} A_{11} & \cdots & A_{1m} \\ \vdots & & \vdots \\ A_{m1} & \cdots & A_{mm} \end{pmatrix}$$

l'auteur définit la partition suivante

$$D = \begin{pmatrix} A_{11} & 0 & \cdots & 0 \\ \vdots & & & \vdots \\ A_{m1} & \cdots & & A_{mm} \end{pmatrix}, \quad C = -\begin{pmatrix} 0 & A_{12} & \cdots & A_{1m} \\ \vdots & & & \vdots \\ 0 & \cdots & & 0 \end{pmatrix}.$$

Il établit une condition de convergence analogue à celle de Gauss-Siedel, c'est-à-dire

$$\rho\,(B^{-1}\,C) < 1 \qquad (3.20)$$

$\rho$ représente la norme spectrale (i.e. la plus grande valeur propre de la matrice).

Ortega et Rheinboldt (70) mentionnent la possibilité d'étendre la méthode de Gauss-Siedel par bloc à un système non-linéaire et énoncent une condition de convergence semblable portant sur les parties de de la matrice du Jacobien, Toutefois les preuves ont été laissées en exercice par les auteurs.

Il n'a pas été possible de vérifier directement la condition de convergence à cause de sa complexité. Nous procédons de façon indirecte en transformant le système (3.19) de façon à obtenir une condition de convergence plus simple qui pourra être vérifiée.

La transformation consiste à éliminer par substitution les variables $k_q$ et ne conserver qu'une seule équation et une seule variable par marché. Utilisons (2.19) pour exprimer $k_q$ en fonction de $u_r$

$$k_q = u_r \; A_q - P_q$$

En substituant dans $\sum_{q \in Q_r} k_q = d^r$ nous obtenons

$$\sum_{q \in Q_r} u_r \; A_q - \sum_{q \in Q_r} P_q = d^r$$

Donc $$u_r - \sum_{q \in Q_r} P_q = d_r$$

Rappelons qu'à l'équation (2.11) $P_q$ avait initialement été défini en fonction des passagers qui tentent d'utiliser une route. C'est sous cette forme (i.e. en fonction des $a_q$) qu'il est utilisé ici. Puisque par (2.18) $a_q = u_r A_q$ les $P_q$ s'expriment directement comme fonctions des $u_r$ et le système (3.19) devient

$$u_r - \sum_{q \in Q_r} P_q(u) = d^r \quad (\forall r) \; . \tag{3.21}$$

La méthode de Gauss-Siedel simple pour les systèmes non-linéaires s'applique pour résoudre le système (3.21) et correspond à la méthode de décomposition discutée précédemment pour résoudre (3.19) et (3.2). Elle se distingue seulement au niveau de la résolution des sous-problèmes qui sont remplacés ici par la résolution d'une équation non-linéaire. La résolution d'une telle équation non-linéaire est probablement plus simple que la résolution des sous-problèmes précédents, mais en pratique nous préférons résoudre les sous-problèmes en utilisant l'algorithme développé pour traiter les modèles P1 et P2.

En vue d'étudier la convergence de la méthode de Gauss-Siedel pour le système (3.21) évaluons son Jacobien.

Soit $\Gamma_r(u) = u_r - \sum_{q \in Q_r} P_q(u) - d^r = 0$ .

Par (2.11)

$$\Gamma_r(u) = u_r - \sum_{q \in Q_r} u_r A_q \sum_{a \in F_q} \frac{E_a(x_a)}{x_a} - d^r = 0$$

avec $x_a = \sum_{p \in E_a} a_p = \sum_{p \in E_a} u_{r_p} A_p$ .

Le Jacobien s'exprime ainsi

$$\frac{\partial \Gamma_r(u)}{\partial u_r} = 1 - \sum_{q \in Q_r} A_q \sum_{a \in F_q} \frac{E_a(x_a)}{x_a} - u_r \sum_{q \in Q_r} A_q^2 \sum_{a \in F_q} \frac{1}{x_a} \left[ E_a'(x_a) - \frac{E_a(x_a)}{x_a} \right]$$

$$\frac{\partial \Gamma_r(u)}{\partial u_s} = - u_r \sum_{q \in Q_r} A_q \sum_{a \in F_q} \frac{1}{x_a} \left[ E_a'(x_a) - \frac{E_a(x_a)}{x_a} \right] \sum_{p \in E_a \cap Q_s} A_p \qquad (3.22) .$$

Pour énoncer la condition de convergence, il faut partitionner le Jacobien $J(u)$ en $B(u)$ et $C(u)$

$$B(u) = \begin{pmatrix} J_{11} & 0 & \cdots & 0 \\ \vdots & & & \vdots \\ \vdots & & & 0 \\ J_{m1} & & & J_{mm} \end{pmatrix} \quad , \quad C(u) = - \begin{pmatrix} 0 & J_{12} & \cdots & J_{1m} \\ \vdots & & & \vdots \\ \vdots & & & J_{m-1,m} \\ 0 & & & 0 \end{pmatrix}$$

Le résultat suivant est démontré par Ortega et Rheinboldt (70). Si $u^*$ est la solution optimale et $u^k$ la solution à l'itération k, alors

$$|u^k - u^*| \leq |u^0 - u^*| \quad \rho\left(B^{-1}(u)\, C(u)\right)^k$$

et évidemment $\{u^k\} \to u^*$ si $\rho\left(B^{-1}(u)\, C(u)\right) < 1$ (3.23)

($\rho$ représente la norme spectrale)

Cette condition est difficile à vérifier mais les deux théorèmes suivants dus aux mêmes auteurs permettent d'obtenir une condition suffisante plus facilement vérifiable.

Théorème: Si $J = B - C$, $C \geq 0$, $B^{-1} \geq 0$ et

$$J^{-1} \geq 0 \text{ alors } \rho(B^{-1} C) < 1 \qquad (3.24)$$

Théorème: Si $A_{ii} > 0$, $A_{ij} \leq 0$ si $i \neq j$ et

$$(\forall i) \quad A_{ii} \geq \sum_j |A_{ij}|$$

$$\text{alors } A^{-1} \text{ existe et } A^{-1} \geq 0 \qquad (3.25)$$

En vue d'appliquer le théorème (3.24), montrons premièrement que $C \geq 0$

Etudions d'abord le terme $\left[E_a'(x_a) - \frac{E_a(x_a)}{x_a}\right]$. On sait que $E_a(0) = 0$ et que $E_a(x_a)$ est une fonction convexe. Donc par l'inégalité du gradient

$0 = E_a(0) \geq E_a(x_a) - E_a'(x_a)\ x_a$ et par conséquent

$$\left[E_a'(x_a) - \frac{E_a(x_a)}{x_a}\right] \geq 0$$

Il en découle que $\frac{\partial r_r(u)}{\partial u_s} \leq 0$ car il est constitué d'un ensemble de facteurs positifs précédés du signe négatif (3.22).

Donc $C \geq 0$

Les conditions $B^{-1} \geq 0$ et $J^{-1} \geq 0$ sont démontrées à l'aide du théorème (3.25). Regardons d'abord $J^{-1}$

Il est évident que $J_{ij} \leq 0$ si $i \neq j$ puisque le Jacobien est symétrique et que sa partie au-dessus de la diagonale est égale à $-C \leq 0$

Vérifions simultanément que $J_{ii} > 0$ et que $J_{ii} \geq \sum_j |J_{ij}|$ en démontrant que $J_{ii} + \sum_j J_{ij} > 0$ .

$$\frac{\partial \Gamma_r(u)}{\partial u_r} - \sum_{s \neq r} \frac{\partial \Gamma_r(u)}{\partial u_s} =$$

$$1 - \sum_{q \in Q_r} A_q \sum_{a \in F_q} \frac{E_a(x_a)}{x_a} - u_r \sum_{q \in Q_r} A_q \sum_{a \in F_q} \frac{1}{x_a} \left[ E_a'(x_a) - \frac{E_a(x_a)}{x_a} \right] \sum_{p \in E_a} A_p$$

Car $A_q + \sum_{s \neq r} \sum_{p \in E_a \cap Q_s} A_p = \sum_{p \in E_a \cap Q_r} A_p + \sum_{s \neq r} \sum_{p \in E_a \cap Q_s} A_p$

$$= \sum_s \sum_{p \in E_a \cap Q_s} A_p = \sum_{p \in E_a} A_p$$

En utilisant $A_p = \frac{a_p}{u_{s_p}}$ et sa conséquence

$$u_r \sum_{p \in E_a} A_p = \sum_{p \in E_a} a_p \frac{u_r}{u_{s_p}} \leq x_a \operatorname{Max}_s \frac{u_r}{u_s}$$

cette expression peut être bornée inférieurement par

$$1 - \sum_{q \in Q_r} A_q \sum_{a \in F_q} \frac{E_a(x_a)}{x_a} - \sum_{q \in Q_r} A_q \sum_{a \in F_q} \left[ E_a'(x_a) - \frac{E_a(x_a)}{x_a} \right] \operatorname{Max} \frac{u_r}{u_s} \quad (3.26)$$

Etudions les valeurs possibles pour le dernier terme. En revenant à la définition de $E_a(x_a)$ on obtient facilement que

$$\left[E_a'(x_a) - \frac{E_a(x_a)}{x_a}\right] = \frac{c_a}{x_a}\left[1 - F\left(\frac{c_a - x_a}{\lambda_a\, x_a}\right)\right] \quad *$$

Le tableau suivant présente les valeurs de $\frac{c_a}{x_a}\left[1-F\left(\frac{c_a-x_a}{\lambda_a\, x_a}\right)\right]$ en fonction de $\lambda_a$ et du rapport $x_a/c_a$

Tableau 2 Valeur de $E_a'(x_a) - E_a(x_a)/x_a$

| $x_a/C_a$ | .6 | .7 | .8 | .9 | 1 | 1.1 | 1.2 | 1.4 |
|---|---|---|---|---|---|---|---|---|
| $\lambda_a = .3$ | .03 | .11 | .25 | .38 | .5 | .57 | .6 | .59 |
| $\lambda_a = .4$ | .09 | .23 | .33 | .42 | .5 | .54 | .55 | .54 |

Comme les $x_a$ sont généralement plus petits que $.8c_a$ les

$\left[E_a'(x_a) - \frac{E_a(x_a)}{x_a}\right]$ sont de l'ordre de .1, .2 et .3 . La plupart des routes ne comportant qu'un ou deux arcs, il s'ensuit que la plupart des valeurs de $\sum_{a \in F_q}\left[E_a'(x_a) - \frac{E_a(x_a)}{x_a}\right]$ se situent entre .1 et .5 . De plus, les quelques routes comptant plus de deux arcs, ont des coefficients d'attraction faibles donc dans la plupart des cas

* F est la fonction de répartition de la loi normale.

$$\sum_{q\in Q_r} A_q \sum_{a\in F_q} \left[ E_a'(x_a) - E_a \frac{(x_a)}{x_a} \right] \leq .55$$

D'autre part, les valeurs des $u_s$ ne sont pas très différentes les unes des autres, car notre méthode de division des origines-destinations assure des marchés de taille semblable .

Nous supposerons donc ici que $\frac{u_r}{u_s} \leq 1.5$ . Nous obtenons donc une borne de $.55 * 1.5 = .83$ pour le dernier terme de (3.26).

Etudions maintenant le second terme de (3.26) $\sum_{a\in F_q} \frac{E_a(x_a)}{x_a}$ est la proportion des passagers refusés le long d'une route. Cette proportion ne devrait pas excéder 15% dans un réseau où un service minimal est assuré. Ainsi $\sum_{q\in Q_r} A_q \sum_{a\in F_q} \frac{E_a(x_a)}{x_a} < \sum_{q\in Q_r} A_q .15 = .15$

En substituant ces bornes dans (3.26)

$$\frac{\partial \Gamma_r(u)}{\partial u_r} - \sum_{s\neq r} \frac{\partial \Gamma_r(u)}{\partial u_s} \geq 1 - .15 - .83 > 0 \qquad (3.27)$$

Ceci démontre que dans la plupart des cas la matrice J est à diagonale dominante. Pour plus de sécurité, la condition de diagonale dominante pourrait être vérifiée à chaque itération à partir des valeurs numériques du Jacobien. Si par malheur elle n'était pas vérifiée, il suffirait de revoir la division en marché de façon à mieux équilibrer les $u_r$ .

Dans le cas où la condition (3.27) est satisfaite, le théorème (3.25) peut être appliqué et $J^{-1} \geq 0$ . De même comme B est obtenu de J en remplaçant certains éléments hors diagonale par 0, le résultat (3.27)

assure aussi que B est à diagonale dominante et que $B_{ii} > 0$ . Ainsi par le théorème (3.25) $B^{-1} \geq 0$ .

Ceci complète la vérification des conditions assurant que la méthode de Gauss-Siedel converge pour ce problème.

On peut aussi montrer directement que $B^{-1} \geq 0$ sans aucune hypothèse sur la grandeur des coefficients en utilisant le fait que $(B^{-1})_{ij} = (-1)^{i+j} \frac{\det B^{ij}}{\det B}$ . On vérifie assez facilement que $\det B > 0$ et que signe $(\det B^{ij}) = (-1)^{i+j}$ .

### 3.1.5 Convergence de la méthode de décomposition pour le modèle P4

Tout comme pour le modèle P3, considérons l'utilisation de Gauss-Siedel pour résoudre le problème réduit en exprimant toutes les variables en fonction des $u_r$. Pour obtenir cette formulation du problème P4 à partir de (3.21), replaçons $d^r$ par l'expression de $v^r$ en fonction de $u_r$ donné par la relation (2.32). Ainsi,

$$u_r - \sum_{q \in Q_r} P_q(u) - d^r + \left(\frac{1-\lambda_r}{\lambda_r}\right)(u_r - d^r) = 0 \quad (\forall r) .$$

$$\text{Donc} \quad \bar{\Gamma}_r(u) = u_r + \left(\frac{1-\lambda_r}{\lambda_r}\right)u_r - \sum_{q \in Q_r} P_q(u) + \frac{d_r}{\lambda_r} = 0 \quad (\forall r) . \tag{3.28}$$

Le Jacobien de $\bar{\Gamma}$ s'exprime en fonction de celui de $\Gamma$ de la façon suivante

$$\frac{\partial \bar{\Gamma}_r(u)}{\partial u_r} = \frac{\partial \Gamma_r(u)}{\partial u_r} + \left(\frac{1-\lambda_r}{\lambda_r}\right)$$

$$\frac{\partial \bar{\Gamma}_r(u)}{\partial u_s} = \frac{\partial \Gamma_r(u)}{\partial u_s} \cdot$$

Il en découle facilement que les conditions de convergence que nous avons vérifiées pour $\Gamma$ sont aussi vérifiées par $\bar{\Gamma}$. Il est intéressant de remarquer que la condition de diagonale dominante du Jacobien est encore plus sûrement vérifiée. En effet,

$$\frac{\partial \bar{\Gamma}_r(u)}{\partial u_r} - \sum_{s \neq r} \frac{\partial \bar{\Gamma}_r(u)}{\partial u_r} = \frac{\partial \Gamma_r(u)}{\partial u_r} - \sum_{s \neq r} \frac{\partial \Gamma_r(u)}{\partial u_s} + \left(\frac{1-\lambda_r}{\lambda_r}\right)$$

$$\geq \left(\frac{1-\lambda_r}{\lambda_r}\right) + 1 - .15 - .83$$

Par exemple, si $\lambda_r = .75$, $\left(\frac{.25}{.75}\right) = .33$ la condition reste satisfaite même si $\frac{u_r}{u_s} = 1{,}7$ et $\sum_{q \in Q_r} A_q \sum_{a \in P_q} \left[E'_a(x_a) - \frac{E_a(x_a)}{x_a}\right] = .7$

Il est intéressant de constater que le modèle avec demande variable est plus robuste vis-à-vis la convergence que celui avec demande fixe. **C'est un autre exemple où l'utilisation d'hypothèses dites simplificatrices complique le problème.**

### 3.1.6 Vitesse de convergence de la méthode de décomposition

La méthode de décomposition demande de parcourir plusieurs fois l'ensemble de marchés en réoptimisant les sous-problèmes associés. Nous avons démontré pour les quatre modèles que cette méthode converge

vers une solution optimale. Voyons maintenant combien de parcours des sous-problèmes sont nécessaires avant d'atteindre une solution satisfaisante.

Les problèmes associés aux modèles P1 et P2 sont considérés du point de vue de la programmation mathématique. La convexité de l'objectif garantit la convergence mais il n'existe pas de résultat théorique sur la vitesse de convergence des méthodes de décomposition pour traiter un problème de programmation mathématique. En pratique, la convergence est rapide et très stable. De nombreux essais du modèle P1 ont été réalisés sur l'ensemble du réseau d'Air Canada en variant la demande, les coefficients d'attractions et la constante K. Tous ces essais ont nécessité de 10 à 12 parcours des sous-problèmes pour obtenir des répartitions des passagers ajustées au passager près. Parcourir dix fois les 1,500 sous-problèmes peut demander jusqu'à 15,000 résolutions de sous-problèmes. Ce nombre a pu être réduit à moins de 3,000 en ne résolvant à chaque parcours que les sous-problèmes dont les conditions d'optimalité ne sont pas respectées au-delà d'une certaine tolérance. Cette tolérance est prise élevée au premier parcours et est abaissée à chaque parcours jusqu'à l'obtention de la précision désirée. En utilisant la méthode de décomposition avec cette stratégie, la répartition des passagers dans l'ensemble du réseau d'Air Canada demande vingt secondes d'unité centrale sur un CONTROL DATA 173 (dix secondes sur un C.D.C. 74).

Les problèmes associés aux modèles P3 et P4 sont traités comme des systèmes d'équations. Dans ce cas, la procédure de décomposition coïncide avec la méthode de Gauss-Siedel et le résultat cité en (3.23) donne une mesure théorique de la vitesse de convergence. Toutefois, cette borne théorique ne fournit pas de résultats pratiques à cause des surestimations faites lors de son évaluation. Du point de vue pratique, un premier essai du modèle P3 avec des données très approximatives a nécessité dix parcours des sous-problèmes. Le nombre de parcours a pu être réduit à sept en introduisant un facteur de relaxation.

La relaxation consiste à modifier l'itération (3.4) de la façon suivante.

$$S_r^{\omega}(H) = H - \omega \left(H - S_r(H)\right) \quad 1 \leq \omega < 2$$

Le cas $\omega=1$ correspond à la méthode standard. Le cas $\omega > 1$ correspond à se déplacer suivant la direction allant de H vers $S_r(H)$ jusqu'à un point au-delà de $S_r(H)$.

Si on considère le système (3.21), lors de la résolution on devait définir $u^{k+1}$ à partir de $u^k$ de telle sorte que $u_s^{k+1} = u_s^k$ si $s \neq r$ et que $u_r^{k+1}$ soit la solution de

$$x - \sum_{q \in Q_r} P_q(u_1, \ldots, u_{r-1}, x, u_{r+1}, \ldots, u_m) = d_r$$

L'itération relaxe devient

$$u^{k+1}(\omega) = u^k - \omega(u^k - u^{k+1})$$

Notons de plus que la convergence de la méthode de Gauss-Siedel garantit la convergence de la méthode de relaxation pour $\omega < 2$.

## 3.2 Algorithmes de résolution des sous-problèmes

Chaque sous-problème correspond au choix des routes par les passagers d'un marché étant donné que certaines places sont déjà utilisées par les passagers des autres marchés. Le sous-problème du problème (3.1) prend la forme suivante pour le marché r

$$\text{Min } f(h_{Q_r}, h_{\overline{Q}_r})$$

$$\sum_{q \in Q_r} h_q = d^r$$

$$h_q \geq 0 \quad (\forall q \in Q_r) \tag{3.29}$$

Les notations $h_{Q_r}$ et $h_{\overline{Q}_r}$ ont été définies à la section 3.1. Dans ce cas, $h_{\overline{Q}_r}$ est prédéterminé seul $h_{Q_r}$ est variable.

Les conditions d'optimalité de Kuhn-Tucker pour ce sous-problème sont:

Il existe $u_r$ tel que

$$\frac{\partial f(h_{Q_r}, h_{\overline{Q}_r})}{\partial h_q} \begin{cases} = u_r & \text{si} \quad h_q \geq 0 \\ \geq u_r & \text{si} \quad h_q = 0 \end{cases} \quad (\forall q \in Q_r)$$

$$\sum_{q \in Q_r} h_q = d^r$$

$$h_q \geq 0 \quad (\forall q \in Q_r) \tag{3.30}$$

D'autre part le sous-problème du problème 3.2 pour le marché r présente la structure suivante:

Il existe $u_r$ tel que

$$C_q(k_{Q_r}, k_{\overline{Q}_r}) \begin{cases} = u_r & \text{si} \quad k_q > 0 \\ \geq u_r & \text{si} \quad k_q = 0 \end{cases} \quad (\forall q \in Q_r)$$

$$\sum_{q \in Q_r} k_q = d^r$$

$$k_q \geq 0 \quad (\forall q \in Q_r) \tag{3.31}$$

où $k_{Q_r}$ et $k_{\bar{Q}_r}$ sont définis de façon analogue à $h_{Q_r}$ et $h_{\bar{Q}_r}$ .

Tout comme nous l'avons vu pour les problèmes complets, les conditions d'optimalité (3.30) du sous-problème (3.29) montrent qu'il est un cas particulier du sous-problème (3.31). Nous allons maintenant montrer que le problème (3.31) peut être exprimé sous une forme similaire à celle du problème (3.29). Rappelons que cette réciproque n'était pas vraie dans le cas des problèmes globaux mais elle est vraie dans le cas des sous-problèmes.

Pour montrer l'existence d'un objectif associé au problème (3.3.1), rappelons la structure particulière des sous-problèmes. La définition des routes, fait à la section 2.1.1 en considérant seulement les chaînes utilisables, assure que deux routes d'un même marché n'ont pas d'arc commun. Le nombre de passagers en excès sur une route est donc indépendant des nombres de clients tentant d'utiliser les autres routes du même marché. Ainsi $C_q(k_{Q_r}, k_{\bar{Q}_r})$ ne dépend que de la variable $k_q$ et on peut définir un objectif f:

$$f(k_{Q_r}, k_{\bar{Q}_r}) = \sum_{q \in Q_r} \int_0^{k_q} C_q(x, k_{\bar{Q}_r})\, dx$$

Puisque $\frac{\partial f}{\partial k_q}(k_{Q_r}, k_{\bar{Q}_r}) = C_q(k_q, k_{\bar{Q}_r})$ ,

cet objectif permet de définir un problème d'optimisation analogue à (3.29) dont les conditions d'optimalité de Kuhn-Tucker (3.30) coincident avec (3.31).

L'équivalence entre (3.31) et (3.29) permet de résoudre les sous-problèmes des modèles P1, P2 et P3 en utilisant les mêmes méthodes d'optimisation. Ces méthodes sont toutefois adaptées pour ne pas nécessiter le calcul explicite de l'objectif dans le cas du problème (3.31). La notation sera celle des modèles P1 et P2 (h représentant les variables) même si nous traitons simultanément les sous-problèmes du modèle 3 où les variables sont de nature différente. Les méthodes de résolutions doivent être adaptées pour résoudre les sous-problèmes du modèle P4 qui présente une structure légèrement différente. Ces adaptations seront discutées à la section 3.2.8.

### 3.2.1 Méthode de direction réalisable

Nous avons adapté et évalué plusieurs variantes de la méthode des directions réalisables de Zoutendijk pour la résolution des sous-problèmes. Cette méthode de type primale génère une suite de solutions réalisables qui converge vers une solution optimale. Comme la solution est réalisable à chaque itération, il est possible de s'arrêter quand l'optimalité est jugée satisfaisante.

Chaque solution $h^k$ de la suite est calculée à partir de $h^{k-1}$ son prédécesseur de la façon suivante:

a) Déterminer d une direction de descente réalisable; i.e. il existe un $\epsilon > 0$ tel que si $0 \leq \sigma \leq \epsilon$ alors $h^{k-1} + \sigma d$ est réalisable et $f(h^{k-1} + \sigma d) < f(h^{k-1})$ .

b) Effectuer le déplacement optimal dans cette direction; i.e. $h^k = h^{k-1} + \bar{\sigma} d$ où $\bar{\sigma}$ est tel que $f(h^{k-1}+\bar{\sigma}d) = \underset{\sigma \in R}{\text{Min}}\ f(h^{k-1}+\sigma d)$

où $R = \{\sigma:\ h^{k-1} + \sigma d$ est réalisable$\}$ .

Les différentes variantes considérées se distinguent par le choix de la direction de descente. Le tableau suivant présente les quatre méthodes étudiées et les particularités de la direction correspondante pour un problème ayant une seule contrainte.

Tableau 3 : Caractéristiques des directions de descente des algorithmes étudiés.

| Algorithme | Caractéristique de la direction |
|---|---|
| Simplexe-convexe | Taux de descente maximal quand une seule variable augmente et une seule variable diminue |
| Gradiant réduit | Taux de descente maximal quand une seule variable diminue. |
| Gradiant | Taux de descente maximal |
| Second ordre | Descente maximal en considérant l'approximation du second ordre de l'objectif. |

La qualité de la direction augmente à mesure qu'on descend dans le tableau. D'autre part,nous verrons que l'effort de calcul pour déterminer ces directions est à peu près le même dans les quatre cas.

### 3.2.2 Calcul du pas optimal

Avant d'étudier chacune des directions,montrons que nous pouvons calculer le déplacement optimal sans connaître explicitement la fonction économique. Ce calcul est fait à partir des $\frac{\partial f(h)}{\partial h_q}$

Soit $P(\sigma) = \nabla f\,(h^k + \sigma d)\cdot d$ et $\sigma^0$ la valeur maximale de $\sigma$ tel que $(h^k + \sigma d)$ soit réalisable.

$$\sigma^0 = \min_{q \,\ni d_q<0} \left\{ -\frac{h_q^k}{d_q} \right\}$$

La méthode proposée pour le calcul du déplacement optimal découle directement de l'équivalence entre les problèmes suivants.

Trouver $\bar{\sigma} \in [0, \sigma^0]$ minimisant $f(h^k + \sigma d)$ (3.32)

Si $P(\sigma^0) \geq 0$, prendre $\bar{\sigma}$ égale à la racine de $P(\sigma)$ (3.33)

sinon prendre $\bar{\sigma} = \sigma^0$

Lemme: Si f est convexe les problèmes (3.32) et (3.33) ont la même solution

Preuve: $\dfrac{\partial f(h^k + \sigma d)}{\partial \sigma} = \nabla f(h^k + \sigma d)\cdot d = P(\sigma)$ .

$P(0) < 0$ car d est une direction de descente.

$P(\sigma)$ est croissante car f est convexe.

Ainsi $P(\sigma^0) \geq 0$ si et seulement si $P(\sigma)$ s'annule sur $[0, \sigma^0]$.

Si $P(\sigma)$ croissante s'annule sur $[0, \sigma^0]$,

alors $\bar{\sigma}$ minimise $f(h^k + \sigma d)$ équivaut à $P(\bar{\sigma}) = 0$

car P est croissante.

Si $P(\sigma)$ ne s'annule pas sur $[0, \sigma^0]$,

alors $\bar{\sigma}$ minimise $f(h^k + \sigma d)$ équivaut à $\bar{\sigma} = \sigma^0$

car $P(\sigma)$ est négatif et f est décroissante pour $\sigma \in [0, \sigma^0]$.

Le problème (3.33) est résolu en trouvant la racine de $P(\sigma)$ par bisection de l'intervalle $[0, \sigma^o]$. Au début $P(0)$ et $P(\sigma^o)$ sont de signes différents. Après chaque bisection, nous conservons l'intervalle contenant le changement de signe de P. Ainsi après k évaluation $|\sigma^k - \bar{\sigma}| \leq \sigma^o (\frac{1}{2})^k$. Cette méthode converge plus rapidement que la méthode des sections dorées pour la recherche du minimum de f. Avec la méthode des sections dorées

$$|\sigma^k - \bar{\sigma}| \leq \sigma^o (1/\rho)^k$$

où $\rho$ = 1.618... le nombre d'or.

Ainsi pour des raisons d'efficacité, il est préférable de chercher la racine de P que de tenter de minimiser f même quand f est connue explicitement.

### 3.2.3. Simplexe convexe

Etant donné que cette méthode est la moins efficace, limitons sa discussion aux remarques suivantes. La matrice de la base est toujours une matrice-identité de dimension 1, il n'a donc pas lieu de faire de pivots, ni de calculer l'inverse de la base. Un changement de base se résume à la modification d'un pointeur indiquant l'indice de la variable de base. Ainsi, il est avantageux et pas plus coûteux de placer dans la base avant chaque itération, la variable non nulle par rapport à laquelle la dérivée partielle est maximale. Le critère d'entrée choisit d'augmenter la variable par rapport à laquelle la dérivée partielle est minimale et diminue la variable de base. Ainsi, chaque itération augmente une variable et en diminue une autre avec le plus grand taux de diminution de l'objectif, ce taux étant donné par la différence entre la dérivée partielle maximum et la dérivée partielle minimum.

### 3.2.4 Gradiant réduit

Dans ce cas aussi la base est toujours la matrice-identité de dimension 1 et le changement de base se fait sans calcul. Il y a donc encore avantage à placer dans la base avant chaque itération, la variable non nulle par rapport à laquelle la dérivée partielle est maximale. Dénotons par $q_k$ la variable de base de l'itération k. Pour chaque variable hors base $h_q$, le coût marginal $z_q$ est défini comme suit:

$$z_q = \frac{\partial f(h^k)}{\partial h_{q_k}} - \frac{\partial f(h^k)}{\partial h_q}$$

Remarquons que si $h_q > 0$ alors $z_q \geq 0$ par le choix de la base. Ainsi la direction définie de la façon suivante est réalisable:

$$d_q = z_q \quad \text{si} \quad q \neq q_k \text{ , sauf si } h_q = 0 \text{ et } z_q < 0 \text{ alors } d_q = 0.$$

$$d_{q_k} = - \sum_{q \in Q_r - \{q_k\}} d_q \qquad (3.34)$$

Dans cette direction, seul la variable de base diminue. C'est parmi les variables pouvant être diminuées, celle de dérivé maximum. Les autres variables sont augmentées en proportion du gradiant réduit. Cette direction donne le plus grand taux de diminution de l'objectif étant donné qu'une seule variable diminue.

### 3.2.5 Méthode du gradiant

Cette méthode s'apparente à la méthode du même nom pour l'optimisation sans contrainte. Dans ce cas, nous profitons du fait que les contraintes sont relativement simples pour déterminer une direction réalisable se rapprochant le plus possible de celle du gradiant. Cette

direction est obtenue analytiquement sans long calcul, et c'est ce qui fait l'intérêt de la méthode.

Pour réduire la notation prenons

$$C_q = C_q(h_q^k) = \frac{\partial f}{\partial h_q}(h_{Q_r}^k, h_{\bar{Q}_r})$$

$$J_k = \{ q \mid h_q^k = 0 \}$$

La direction recherchée est la solution du problème d'optimisation suivant:

$$\text{Min} \sum_{q \in Q_r} C_q \; d_q$$

$$\sum_{q \in Q_r} d_q = 0$$

$$\sum_{q \in Q_r} d_q^2 \leq 1$$

$$d_q \geq 0 \qquad (\forall q \in J_k) \qquad (3.35)$$

Ce problème définit d comme la direction réalisable qui maximise le taux de descente. Une solution analytique du problème (3.35) peut être déterminée avec une approche semblable à celle qu'on retrouve dans Coheen (77) et Rockafellar (70). Cette méthode analyse les conditions de Kuhn-Tucker du problème (3.35)

K.-T.1 $\quad C_q - u - 2\lambda d_q = 0 \qquad (\forall q \notin J_k)$

K.-T.2 $\quad C_q - u - 2\lambda d_q \geq 0 \qquad (\forall q \in J_k)$

K.-T.3 $(C_q - u - 2\lambda d_q)\, d_q = 0 \qquad (\forall q \in J_k)$

K.-T.4 $\sum_{q \in Q_r} d_q = 0$

K.-T.5 $\sum_{q \in Q_r} d_q^2 \leq 1$

K.-T.6 $\lambda\left( \sum_{q \in Q_r} d_q^2 - 1 \right) = 0$

K.-T.7 $d_q \geq 0 \qquad (\forall q \in J_k)$

K.-T.8 $\lambda \leq 0$ . (3.36)

Montrons d'abord que $\lambda < 0$

$$0 = \sum_{q \in Q_r} (C_q - u - 2\lambda d_q)\, d_q \qquad \text{(par K.-T. 1 et 3)}$$

$$= \sum_{q \in Q_r} C_q d_q - u \sum_{q \in Q_r} d_q - 2\lambda \sum_{q \in Q_r} d_q^2$$

$$\geq \sum_{q \in Q_r} C_q d_q - u\, 0 - 2\lambda \qquad \text{(par K.-T. 4 et 5) .}$$

D'où $\lambda \leq \frac{1}{2} \sum_{q \in Q_r} C_q d_q < 0$ , car d est une direction de descente.

**Exprimons maintenant les $d_q$ en fonction de u et $\lambda$**

Si $q \notin J_k$ , alors

$$d_q = \frac{C_q - u}{2\lambda} \qquad \text{(par K.-T. 1).}$$

Si $q \in J_k$ et $C_q < u$

$C_q - u < 0$ implique $2\lambda d_q < 0$ (par K.-T. 2) .

D'où $d_q \neq 0$ et $(C_q - u - 2\lambda d_q) = 0$ (par K.-T. 3)

et ainsi $d_q = \frac{C_q - u}{2\lambda}$ .

Si $q \in J_k$ et $C_q \geq u$

$C_q - u \geq 0$ implique $(C_q - u)\, d_q \geq 0$ (par K.-T. 7)

et puisque $(C_q - u)\, d_q = 2\lambda\, d_q^2$ (par K.-T. 3)

alors $2\lambda\, d_q^2 \geq 0$ et $d_q^2 \leq 0$ car $\lambda < 0$

donc $d_q = 0$ .

On peut résumer ces résultats comme suit

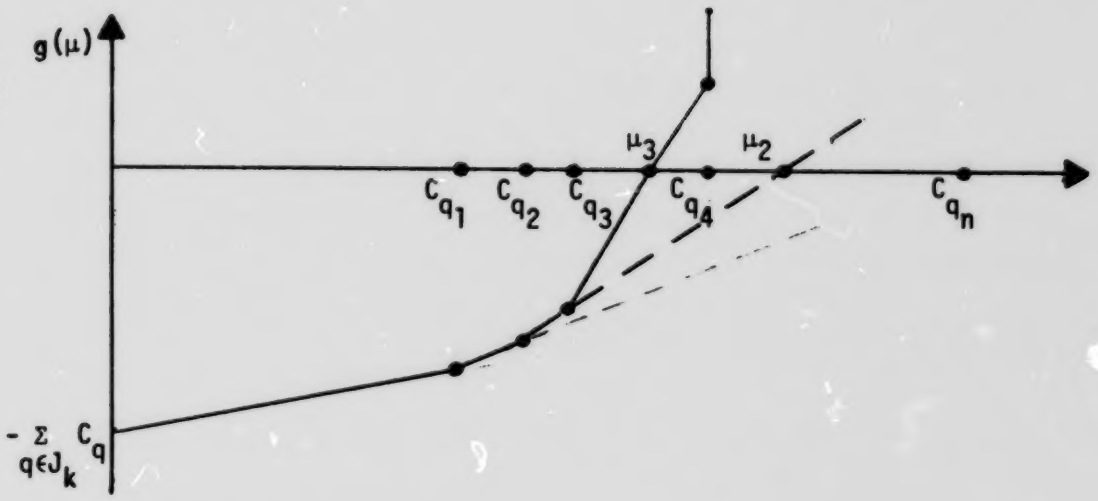
$$d_q(u, \lambda) = \begin{cases} \frac{C_q - u}{2\lambda} & \text{si } q \notin J_k \text{ ou si } q \in J_k \text{ et } C_q < u \ . \\ 0 & \text{si } q \in J_k \text{ et } C_q \geq u \ . \end{cases} \qquad (3.37)$$

Déterminons maintenant la valeur de u pour que $\sum_{q \in Q_r} d_q(u, \lambda) = 0$.

Il est facile de voir que cette valeur de u est indépendante de $\lambda$. Ainsi, considérons $g(u) = \sum_{q \in Q_r} d_q(u, -\frac{1}{2})$ et trouvons sa racine. $g(u)$ est une fonction linéaire par morceau dont les points critiques sont les $u = C_q$ pour $q \in J_k$. Numérotons les éléments de $J_k = \{q_1, \ldots, q_n\}$ tel que $C_{q_1} \leq C_{q_2} \leq \ldots \leq C_{q_n}$ .

FIGURE 13

GRAPHIQUE DE $g(\mu)$ DE LA DIRECTION DU GRADIENT




Pour déterminer $\bar{C}$ une racine de g déterminons pour chaque morceau de g le point $u_j$, j = 0, ..., n qui est racine du prolongement linéaire de ce morceau et ainsi le point satisfaisant la relation $C_{q_j} \leq u_j < C_{q_{j+1}}$ est la racine de g. Les cas frontières j = 0 et j = n demandent seulement $u_0 \leq C_{q_1}$ et $C_{q_n} \leq u_n$ .

En utilisant l'équation du $j^{ième}$ morceau de g nous obtenons

$$u_j = C_{q_j} - g(C_{q_j}) \frac{C_{q_{j+1}} - C_{q_j}}{g(C_{q_{j+1}}) - g(C_{q_j})} \quad . \tag{3.38}$$

L'évaluation du membre de droite de (3.38) donne

$$u_j = \frac{\sum_{q \notin J_k} C_q + \sum_{i=1}^{j} C_{q_i}}{\|Q_r\| - \|J_k\| + j} \qquad (3.39)$$

où $\|A\|$ indique le nombre d'éléments de A.

On peut facilement vérifier que (3.39) tient aussi pour les cas frontières $j = 0$ et $j = n$. Ainsi, $\bar{C}$ est facilement calculé en évaluant séquentiellement $u_0$, $u_1$, ... jusqu'à ce que $C_{q_j} \leq u_j < C_{q_{j+1}}$. Appelons

$$P_k = \{q | q \notin J_k\} \cup \{q_1, \ldots, q_j\}$$

La valeur de $\lambda$ doit être telle que $\sum_{q \in Q_r} d^2_q = 1$. En pratique, il est suffisant de connaître d à une constante près car la longueur du déplacement qui est calculée par la suite est ajustée en conséquence. Ainsi, en prenant $\lambda = -\frac{1}{2}$ et $u = \bar{C}$ nous obtenons de (3.37)

$$d_q = \begin{cases} \bar{C} - C_q & \text{si} \quad q \in P_k \\ 0 & \text{sinon} \end{cases} \qquad (3.40)$$

où $P_k = \{q | q \notin J_k\} \cup \{q_1, \ldots, q_j\}$

Mentionnons que cette direction de descente peut être vue comme la variante de la méthode du gradiant projeté (Luenberger, 73) qui choisit à chaque itération le sous-espace sur lequel est projeté le gradiant de façon à maximiser le taux de descente. Ce sous-espace qui est de dimension $\|P_k\| - 1$ est l'intersection de l'espace des variables dont les indices appartiennent à $P_k$ avec l'hyperplan $\sum_{q \in Q_r} h_q = 0$

### 3.2.6 Méthode du second ordre

Cette méthode utilise comme direction de descente celle qui donne la descente maximale de l'approximation du second ordre de l'objectif. Cette direction peut être calculée,aussi facilement que les précédentes,à cause des particularités des problèmes considérés.

Précisons les particularités du problème traité et les notations utilisées

$$C_q(h_q) = \frac{\partial f}{\partial h_q}(h_{Q_r}, h_{\bar{Q}_r}) > 0 \qquad (3.41)$$

$$S_q(h_q) = \frac{\partial}{\partial h_q} C_q(h_q) > 0 \quad \text{et} \quad \frac{\partial}{\partial h_p} C_q(h_q) = 0 \quad \text{si} \quad p \neq q \qquad (3.42)$$

$$T_q(h_q) = \frac{\partial}{\partial h_q} S_q(h_q) > 0 \quad \text{et} \quad \frac{\partial}{\partial h_p} S_q(h_q) = 0 \quad \text{si} \quad p \neq q \qquad (3.43)$$

Rappelons que $\frac{\partial f}{\partial h_q}(h_{Q_r}, h_{\bar{Q}_r})$ est indépendant de $h_p$ pour tout $p \in Q_r - \{q\}$ puisque les routes d'un même marché n'ont pas d'arcs en commun. $C_q$ est donc fonction de la seule variable $h_q$ et il en est de même pour ses dérivés $S_q$ et $T_q$. La positivité des dérivés est facilement vérifiable.

Soit $h^k = (h_q^k \mid q \in Q_r)$ la solution à l'itération k. Pour définir la direction du second ordre à partir de $h^k$ introduisons les simplifications de notation suivantes $C_q = C_q(h_q^k)$, $S_q = S_q(h_q^k)$, $T_q = T_q(h_q^k)$ et $J_k = \{q \mid h_q^k = 0\}$ . La direction est la solution du problème d'optimisation suivant.

$$\text{Min} \sum_{q \in Q_r} \frac{1}{2} S_q d_q^2 + C_q d_q$$

$$\sum_{q \in Q_r} d_q = 0$$

$$d_q \geq 0 \qquad (\forall q \in J_k) . \tag{3.44}$$

Trouvons analytiquement la solution du problème (3.44) en utilisant les conditions d'optimalité de Kuhn-Tucker.

K.-T.1 $S_q d_q + C_q - u = 0 \qquad (\forall q \notin J_k)$ .

K.-T.2 $S_q d_q + C_q - u \geq 0 \qquad (\forall q \in J_k)$ .

K.-T.3 $(S_q d_q + C_q - u) d_q = 0 \quad (\forall q \in J_k)$ .

K.-T.4 $\sum_{q \in Q_r} d_q = 0$ .

K.-T.5 $d_q \geq 0 \qquad (\forall q \in J_k)$ . (3.45)

Pour résoudre ce système, procédons comme à la section précédente. Exprimons $d_q$ comme fonction de u.

Si $q \notin J_k$ alors (par K.-T.1) $d_q = \frac{u - C_q}{S_q}$ .

Si $q \in J_k$ et $u - C_q > 0$, alors (par K.-T.2) $S_q d_q = u - C_q > 0$

et ainsi $d_q > 0$ car $S_q > 0$ . Par K.-T.3

$$S_q d_q + C_q - u = 0 \text{ et alors } d_q = \frac{u - C_q}{S_q} .$$

Si $q \in J_k$ et $u-C_q \leq 0$ alors $(C_q-u)\, d_q \geq 0$ et par K.-T.3

$S_q\, d_q^2 \leq 0$ . Donc $d_q = 0$ car $S_q > 0$ .

Ces résultats se résument ainsi

$$d_q(u) = \begin{cases} \dfrac{u-C_q}{S_q} & \text{si } q \notin J_k \text{ et si } q \in J_k \text{ et } u-C_q > 0 . \\ 0 & \text{si } q \in J_k \text{ et } u-C_q \leq 0 . \end{cases} \qquad (3.46)$$

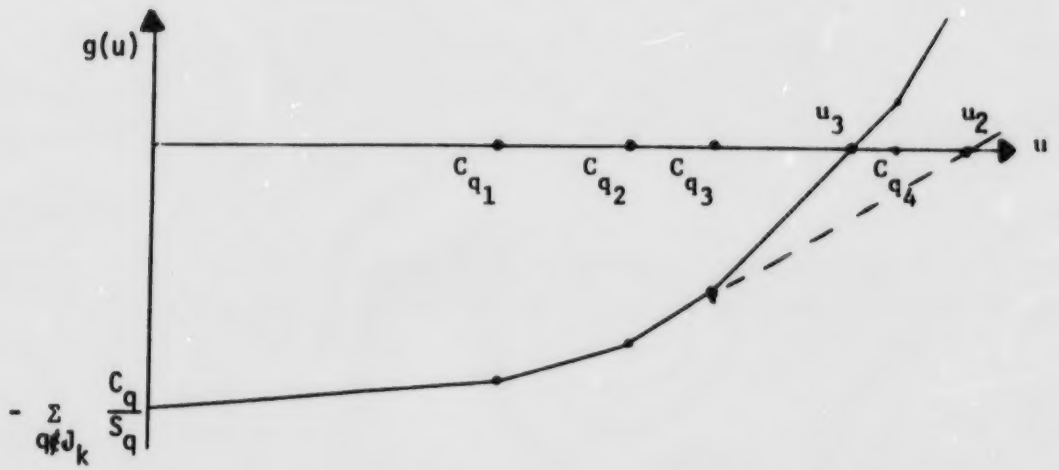
Déterminons maintenant la valeur $\bar{C}$ pour que

$$g(u) = \sum_{q \in Q_r} d_q(u) = 0 .$$

Tout comme précédemment, numérotons les $C_q$, $q \in J_k$ tel que $C_{q_1} \leq C_{q_2} \leq \ldots \leq C_{q_n}$.

FIGURE 14

GRAPHIQUE DE g(u) DE LA DIRECTION DU SECOND ORDRE



Calculons le point où s'annule le prolongement linéaire de chaque morceau:

$$u_j = C_{q_j} - g(C_{q_j}) \frac{C_{q_{j+1}} - C_{q_j}}{g(C_{q_{j+1}}) - g(C_{q_j})} \tag{3.47}$$

et en évaluant le membre de droite

$$u_j = \frac{\sum_{q \notin J_k} \frac{C_q}{S_q} + \sum_{i=1}^{j} \frac{C_{q_i}}{S_{q_i}}}{\sum_{q \notin J_k} \frac{1}{S_q} + \sum_{i=1}^{j} \frac{1}{S_{q_i}}} \tag{3.48}$$

Ainsi $\bar{C}$ est facilement déterminé en évaluant séquentiellement $u_0, u_1, u_2, \ldots$ jusqu'à ce que

$$C_{q_j} \leq \bar{C} = u_j < C_{q_{j+1}} \tag{3.49}$$

Tout comme précédemment, dénotons $P_k = \{q | q \notin J_k\} \cup \{q_1, \ldots, q_j\}$

La direction d définie par (3.44) est donc

$$d_q = \begin{cases} \dfrac{\bar{C} - C_q}{S_q} & \text{si} \quad q \in P_k \\ 0 & \text{sinon} \end{cases} \tag{3.50}$$

Remarquons que $d_q$ et $\bar{C}$ sont notés $d_q^k$ et $\bar{C}^k$ quand il est utile de préciser l'itération.

Maintenant nous allons montrer qu'un algorithme utilisant cette direction converge vers la solution optimale. Cette preuve ne

peut pas utiliser la méthode de Zangwill car la direction du second ordre n'est pas une application fermée. Nous donnons une preuve directe montrant que la convergence est d'ordre deux.

Pour obtenir ce résultat, il faut modifier le pas utilisé par l'algorithme. Plutôt que de prendre $h^{k+1} = h^k + \sigma^k d^k$ qui minimise l'objectif dans cette direction, nous définissons le pas $\sigma^k$ en se référant aux conditions d'optimalité de Kuhn-Tucker (3.30). Soit $\sigma^o > 0$ la plus grande valeur que peut prendre $\sigma \in (0,1]$ tel que $h^k + \sigma^o d^k$ reste réalisable. Alors $\sigma^k$ est défini comme étant le $\sigma \in (0, \sigma^o]$ qui minimise

$$E_{k+1} = \sum_{q \in P_k} \left[ C_q(h_q^k + \sigma d_q^k) - u(\sigma) \right] \qquad (3.51)$$

où $u(\sigma) = \min_{q \in Q_r} C_q(h_q^k + \sigma d_q^k)$ .

Etant donné que $P_k = \{q \mid h_q^{k+1} = h_q^k + \sigma^k d_q^k > 0\}$ en choisissant $\sigma^k$ ainsi, on détermine un $h^{k+1}$ de façon à rendre minimal le total des écarts dans la relation non encore satisfaite de (3.30). Il s'agit de la relation $\left(C_q(h^{k+1}) = u \quad \text{si} \quad h_q^{k+1} > 0\right)$ .

La vitesse de convergence est définie en terme de $E_k$ (3.51). La preuve est divisée en deux parties. Premièrement, il est montré qu'après un nombre fini d'itérations, toutes les variables positives demeurent positives par la suite. Deuxièmement, nous montrons qu'après ce nombre fini d'itération, la convergence de $E_k$ vers zéro est d'ordre 2.

La première partie de la preuve est basée sur les trois lemmes suivants.

Lemme 1 $d_q^k \le 0$ implique $C_q(h_q^{k+1}) \ge \bar{C}^k$ .

Preuve. Rappelons que $h_q^{k+1} = h_q^k + \sigma^k d_q^k$ avec $0 < \sigma^k \le 1$ .

Considérons d'abord le cas où $d_q^k < 0$ .

Par le théorème de Taylor

$$C_q(h_q^{k+1}) = C_q(h_q^k) + \sigma^k d_q^k S_q(h_q^k) + (\sigma^k d_q^k)^2 T_q(\xi)$$

pour $h_q^k \le \xi \le h_q^{k+1}$ .

Alors $C_q(h_q^{k+1}) \ge C_q(h_q^k) + \sigma^k d_q^k S_q(h_q^k)$, car $T_q(\xi) > 0$

$$\ge C_q(h_q^k) + d_q^k S_q(h_q^k), \quad \text{car } \sigma^k \le 1$$

$$\ge C_q(h_q^k) + \bar{C}^k - C_q(h_q^k) \quad \text{par la définition de } d_q^k \quad (3.50)$$

Donc $C_q(h_q^{k+1}) \ge \bar{C}^k$ .

Dans le cas où $d_q^k = 0$ , $C_q(h_q^{k+1}) = C_q(h_q^k)$

et $\bar{C}^k - C_q(h_q^k) \le 0$ par (3.46) .

Donc on a aussi $C_q(h_q^{k+1}) \ge \bar{C}^k$ .

Lemme 2 Si $\bar{C}^k > C_q(0)$ , alors $h_q^{k+1} > 0$ .

Preuve. $h_q^{k+1} = h_q^k + \sigma^k d_q^k$ avec $h_q^k \geq 0$ et $0 < \sigma \leq 1$ .

Si $d_q^k > 0$ alors nécessairement $h_q^{k+1} > 0$ .

Si $d_q^k \leq 0$, alors $C_q(h_q^{k+1}) \geq \overline{C}^k$ par le lemme 1, et

ainsi $C_q(h_q^{k+1}) > C_q(0)$ puisque par hypothèse $\overline{C}^k > C_q(0)$ .

Donc $h_q^{k+1} > 0$ car $C_q(\ )$ est une fonction croissante $\left(\frac{\partial C_q(\ )}{\partial h_q} > 0\right)$ .

Lemme 3. Si les variables positives avant l'itération k demeurent positives à la fin de l'itération k (i.e. si $h_q^k > 0$ implique $h_q^{k+1} > 0$) alors $\overline{C}^{k+1} \geq \overline{C}^k$ .

Preuve. Rappelons que $P_k = \{ q \mid h_q^k > 0\} \cup \{q \mid h_q^k = 0 \text{ et } C_q(h_q^k) < \overline{C}^k\}$

est l'ensemble des indices des variables qui peuvent être modifiées durant l'itération k .

Par définition de d, si

$$h_q^k = 0 \text{ et } C_q(h_q^k) < \overline{C}^k, \text{ alors } d_q^k = \frac{\overline{C}_q - C_q(h_q^k)}{S_q(h_q^k)} > 0$$

et ainsi $h_q^{k+1} = h_q^k + \sigma^k d_q^k > 0$ .

En utilisant l'hypothèse du lemme

$$P_k = \{q \mid h_q^k > 0\} \cup \{q \mid h_q^k = 0 \text{ et } C_q(h_q^k) < \overline{C}\} = \{q \mid h_q^{k+1} > 0\} \subseteq P_{k+1} .$$

**Définissons:**

$$\pi = \sum_{q\in P_k} \frac{C_q(h_q^{k+1})}{S_q(h_q^{k+1})} \Big/ \sum_{q\in P_k} \frac{1}{S_q(h_q^{k+1})} \quad \text{et montrons que} \quad \pi - \bar{C}^k \geq 0. \quad \text{Or}$$

$$\pi - \bar{C}^k = \sum_{q\in P_k} \frac{C_q(h_q^{k+1}) - \bar{C}^k}{S_q(h_q^{k+1})} \Big/ \sum_{q\in P_k} \frac{1}{S_q(h_q^{k+1})}$$

et puisque le dénominateur est positif il suffit de montrer que le numérateur est aussi positif.

Partitionnons $P_k = J_1 \cup J_2 \cup J_3$

où $J_1 = \{p \in P_k \mid C_q(h_q^{k+1}) - \bar{C}^k < 0\}$

$J_2 = \{p \in P_k \mid C_q(h_q^{k+1}) - \bar{C}^k \geq 0 \text{ et } d_q^k < 0\}$

$J_3 = \{p \in P_k \mid C_q(h_q^{k+1}) - \bar{C}^k \geq 0 \text{ et } d_q^k \geq 0\}$

Si $q \in J_1$, alors $C_q(h_q^{k+1}) - \bar{C}^k < 0$ et la contraposé du lemme 1 implique $d_q^k > 0$ .

Ainsi $h_q^{k+1} = h_q^k + \sigma^k d_q^k > h_q^k$, et alors $S_q(h_q^{k+1}) > S_q(h_q^k)$ puisque $S_q(\ )$ est une fonction croissante.

Donc $$\sum_{q\in J_1} \frac{C_q(h_q^{k+1}) - \bar{C}^k}{S_q(h_q^{k+1})} \geq \sum_{q\in J_1} \frac{C_q(h_q^{k+1}) - \bar{C}^k}{S_q(h_q^k)} \quad . \qquad (3.52)$$

Si $q \in J_2$, alors $d_q^k < 0$ ce qui implique $h_q^{k+1} = h_q^k + \sigma^k h_q^k < h_q^k$, et par conséquent $S_q(h_q^{k+1}) < S_q(h_q^k)$ puisque $S_q(\ )$ est une fonction croissante.

$$\text{Donc} \quad \sum_{q\in J_2} \frac{C_q(h_q^{k+1}) - \bar{C}^k}{S_q(h_q^{k+1})} \geq \sum_{q\in J_2} \frac{C_q(h^{k+1}) - \bar{C}^k}{S_q(h_q^k)} \quad . \tag{3.53}$$

Si $q \in J_1 \cup J_2$, alors $d_q^k \neq 0$ .

Par le théorème de Taylor

$$C_q(h_q^{k+1}) - \bar{C}^k = C_q(h_q^k + \sigma^k d_q^k) - \bar{C}^k$$

$$= C_q(h_q^k) + \sigma^k d_q^k \ S_q(h_q^k) + (\sigma^k d_q^k)^2 \ T_q(\xi) - \bar{C}^k$$

$$\geq C_q(h_q^k) + \sigma^k d_q^k \ S_q(h_q^k) - \bar{C}^k \quad \text{car} \quad T_q(\xi) > 0$$

$$\geq (\sigma^k - 1) \ d_q^k \ S_q(h_q^k) \quad \text{par le cas} \quad d_q^k \neq 0 \quad \text{de (3.50)} \ .$$

En subtituant cette dernière relation dans (3.52) et (3.53), nous obtenons

$$\sum_{q\in J_1 \cup J_2} \frac{C_q(h_q^{k+1}) - \bar{C}^k}{S_q(h_q^{k+1})} \geq (\sigma^k - 1) \sum_{q\in J_1 \cup J_2} d_q^k \ .$$

Si $q \in J_3$, alors $C_q(h_q^{k+1}) - \bar{C}^k \geq 0$ et $d_q^k \geq 0$ et ainsi

$$\sum_{q\in J_3} \frac{C_q(h_q^{k+1}) - \bar{C}^k}{S_q(h_q^{k+1})} \geq 0 \geq (\sigma^k - 1) \sum_{q\in J_3} d_q^k \quad \text{car} \quad \sigma^k \leq 1 \ .$$

En regroupant ces deux dernières relations, on obtient

$$\sum_{q\in P_k} \frac{C_q(h_q^{k+1}) - \bar{C}^k}{S_q(h_q^{k+1})} \geq (\sigma^k - 1) \sum_{q\in P_k} d_q^k = 0$$

et ainsi $\pi - \bar{C}^k \geq 0$ .

Utilisons maintenant ce résultat pour montrer que $\bar{C}^{k+1} - \bar{C}^k \geq 0$. Avant d'évaluer $\bar{C}^{k+1}$, rappelons que si $q \in P_{k+1} - P_k$ alors $h_q^{k+1} = 0$ et ainsi $C_q(0) \geq \bar{C}^k$ par le lemme 2.

$$\bar{C}^{k+1} = \frac{\sum_{q\in P_{k+1}} \frac{C_q(h_q^{k+1})}{S_q(h_q^{k+1})}}{\sum_{q\in P_{k+1}} \frac{1}{S_q(h_q^{k+1})}} = \frac{\sum_{q\in P_{k+1}-P_k} \frac{C_q(0)}{S_q(h_q^{k+1})}}{\sum_{q\in P_{k+1}} \frac{1}{S_q(h_q^{k+1})}} + \frac{\sum_{q\in P_k} \frac{C_q(h_q^{k+1})}{S_q(h_q^{k+1})}}{\sum_{q\in P_{k+1}} \frac{1}{S_q(h_q^{k+1})}} .$$

$$\bar{C}^{k+1} \geq \frac{\sum_{q\in P_{k+1}-P_k} \frac{\bar{C}^k}{S_q(h_q^{k+1})}}{\sum_{q\in P_{k+1}} \frac{1}{S_q(h_q^{k+1})}} + \frac{\sum_{q\in P_k} \frac{C_q(h_q^{k+1})}{S_q(h_q^{k+1})}}{\sum_{q\in P_k} \frac{1}{S_q(h_q^{k+1})}} \cdot \frac{\sum_{q\in P_k} \frac{1}{S_q(h_q^{k+1})}}{\sum_{q\in P_{k+1}} \frac{1}{S_q(h_q^{k+1})}} .$$

$$\bar{C}^{k+1} \geq \bar{C}^k \frac{\sum_{q\in P_{k+1}-P_k} \frac{1}{S_q(h_q^{k+1})}}{\sum_{q\in P_{k+1}} \frac{1}{S_q(h_q^{k+1})}} + \pi \frac{\sum_{q\in P_k} \frac{1}{S_q(h_q^{k+1})}}{\sum_{q\in P_{k+1}} \frac{1}{S_q(h_q^{k+1})}} .$$

D'où $\bar{C}^{k+1} \geq \bar{C}^k$ puisque $\pi \geq \bar{C}^k$ .

Ceci complète la preuve du lemme.

Pour montrer qu'après un nombre fini d'itérations toutes les variables positives demeurent positives par la suite, supposons par

contradiction qu'il existe une suite infinie d'itérations où une variable s'annule. Comme il n'y a qu'un nombre fini de variables qui peuvent s'annuler et demeurer nulles, il existe un indice k tel qu'après l'itération k toutes variables qui s'annulent redeviennent positives par la suite. Soit H l'ensemble des indices des variables qui s'annulent après l'itération k. Soit $s \in H$ tel que

$$C_s(0) = \underset{q \in H}{\text{Max}} \; C_q(0) \tag{3.54}$$

Soit $v > k$ une itération où $h_s^v$ devient positive. i.e. $h_s^v = 0$ et $h_s^{v+1} > 0$. Alors $d_s^v > 0$ et $d_s^v = \dfrac{\bar{C}^v - C_s(0)}{S_s(0)} > 0$ implique $\bar{C}^v > C_s(0)$ . La définition de s (3.54) entraîne $\bar{C}^v > C_q(0)$ pour $q \in H$ et ainsi par le lemme 2 il s'ensuit qu'aucune variable de l'ensemble H ne s'annule à l'itération v. Donc aucune variable ne s'annule à l'itération v, car les variables dont l'indice n'appartient pas à H ne peuvent s'annuler à une itération $v > k$.

Comme aucune variable ne s'annule à l'itération v, le lemme 3 implique $\bar{C}^{v+1} \geq \bar{C}^v$ . Ainsi $\bar{C}^{v+1} > C_q(0)$ pour $q \in H$ et l'application alternative des lemmes 2 et 3 entraîne qu'aucune variable ne s'annule par la suite. Ceci contredit l'existence d'une suite infinie d'itérations avec une variable qui s'annule, ce qui complète la première partie de la preuve.

Après un nombre fini d'itérations lorsque toutes les variables positives demeurent positives, établissons une relation entre $E_{k+1}$ et $E_k$ . Cette relation indique que la convergence est d'ordre deux.

A partir de ce moment, les variables nulles qui deviennent positives le demeurent. A cause du nombre fini de variables, il existe donc une itération k. Après laquelle les variables positives demeurent positives et les variables nulles demeurent nulles. Soit $\Gamma$ l'ensemble des indices des variables positives après l'itération k.

Remarquons que $k > k_0$ entraîne $P_k = \Gamma$ puisque

$$P_k = \{ q \mid h_q^k > 0 \} \cup \{ q \mid h_q^k = 0, \bar{C} > C_q \} .$$

En effet, la première composante est $\Gamma$ et la seconde est vide car $h_q^k = o$ et $d_q^k = \dfrac{\bar{C} - C_q}{S_q} > 0$ correspond une variable nulle qui devient positive, ce qui est impossible pour $k > k_0$ .

Remarquons de plus que lorsque $k > k_0$ les composantes $h_q^k$, $q \notin \Gamma$, n'interviennent pas dans la définition de $\bar{C}^k$ et de celle de $d^k$ . Ainsi, pour simplifier la notation, nous pouvons réduire le problème aux variables dont les indices appartiennent à $\Gamma$ .

Soit $0 \leq \sigma \leq 1$ et soit $q_o$ et $q_\sigma$ des indices tels que

$$C_{q_o}(h_{q_o}^k) = \min_{q \in \Gamma} C_q(h_q^k) . *$$

$$C_{q_\sigma}(h_{q_\sigma}^k + \sigma d_{q_o}^k) = \min_{q \in \Gamma} C_q(h_q^k + \sigma d_q^k) .*$$

Par la définition du minimum

$$C_q(h_q^k) - C_{q_o}(h_{q_o}^k) \geq C_q(h_q^k) - C_{q_\sigma}(h_{q_\sigma}^k) . \qquad (3.55)$$

---

* On peut montrer que les minimums pour $q \in \Gamma$ et $q \in Q_r$ coincident.

En utilisant le développement de Taylor pour $q \in I'$ on obtient

$$C_q(h_q^k + \sigma\, d_q^k) = C_q(h_q^k) + \sigma\, d_q^k\ S_q(h_q^k) + \tfrac{1}{2}\sigma^2 (d_q^k)^2\, T_q(\xi_q)\ .$$

En remplaçant $d_q^k$ par sa valeur $\left(\text{i.e. } d_q^k = \dfrac{\bar{c}^k - C_q(h_q^k)}{S_q(h_q^k)}\right)$ on obtient

$$C_q(h_q^k + \sigma\, d_q^k) = C_q(h_q^k) + \sigma\left[\bar{c}^k - C_q(h_q^k)\right] + \tfrac{1}{2}\sigma^2\left[\bar{c}^k - C_q(h_q^k)\right]^2 \frac{T_q(\xi_q)}{S_q(h_q^k)^2}\ .$$

En utilisant cette équation pour $q$ et $q_\sigma$ on obtient

$$C_q(h_q^k + \sigma\, d_q^k) - C_{q_\sigma}(h_{q_\sigma}^k + \sigma\, d_{q_\sigma}^k) =$$

$$(1-\sigma)\left[C_q(h_q^k) - C_{q_\sigma}(h_{q_\sigma}^k)\right] +$$

$$\tfrac{1}{2}\sigma^2\left\{\left[\bar{c}^k - C_q(h_q^k)\right]^2 \frac{T_q(\xi_q)}{S_q(h_q^k)^2} - \left[\bar{c}^k - C_{q_\sigma}(h_{q_\sigma}^k)\right]^2 \frac{T_{q_\sigma}(\xi_{q_\sigma})}{S_{q_\sigma}(h_{q_\sigma}^k)^2}\right\}\ . \qquad (3.56)$$

Soit $Q = \underset{q\in Q_r}{\text{Max}}\ \underset{\xi,h}{\text{Max}}\ \dfrac{T_q(\xi)}{S_q(h)}$ ,

où $\xi$ et $h$ sont dans le domaine réalisable

Comme $T_{q_\sigma}(\xi_{q_\sigma}) > 0$ on peut majorer (3.56) ainsi

$$C_q(h_q^k + \sigma d_q^k) - C_{q_\sigma}(h_{q_\sigma}^k + \sigma d_{q_\sigma}^k) \le (1-\sigma)\left[C_q(h_q^k) - C_{q_\sigma}(h_{q_\sigma}^k)\right] +$$

$$\tfrac{1}{2}\sigma^2\, Q\left[\bar{c}^k - C_q(h_q^k)\right]^2\ . \qquad (3.57)$$

En utilisant la définition de $\bar{C}^k$ donnée en (3.48) et (3.49) et la définition de $E_k$ donnée en (3.51), il est facile de vérifier que

$$\left[\bar{C}^k - C_q(h_q^k)\right]^2 \leq \left[\underset{p\in\Gamma}{\text{Max}}\ C_p(h_p^k) - \underset{p\in\Gamma}{\text{Min}}\ C_p(h_p^k)\right]^2 \leq E_k^2 .$$

En combinant ce résultat avec (3.55) et (3.57) on obtient

$$C_q(h_q^k + \sigma d_q^k) - C_{q_\sigma}(h_{q_\sigma}^k + \sigma\, d_{q_\sigma}^k) \leq (1-\sigma)\left[C_q(h_q^k) - C_{q_o}(h_{q_o}^k)\right] + \tfrac{1}{2}\sigma^2 Q E_k^2 .$$

En prenant le minimum sur $\sigma \in [0,1]$ de la somme sur tous les $q \in \Gamma$ de chaque côté de l'inéquation précédente on obtient

$$E_{k+1} \leq \underset{0\leq\sigma\leq 1}{\text{Min}} \left\{(1-\sigma)\, E_k + \tfrac{1}{2}\sigma^2 \bar{n}\, Q\, E_k^2\right\} \qquad (3.58)$$

où $\bar{n}$ est le nombre de variables positives

L'apparition de $E_{k+1}$ dans cette équation vient du fait qu'aucune variable ne s'annule et par conséquent le $\sigma$ qui minimise la somme pour $q \in \Gamma$ est celui qui définit $h^{k+1}$ et $E_{k+1}$ .

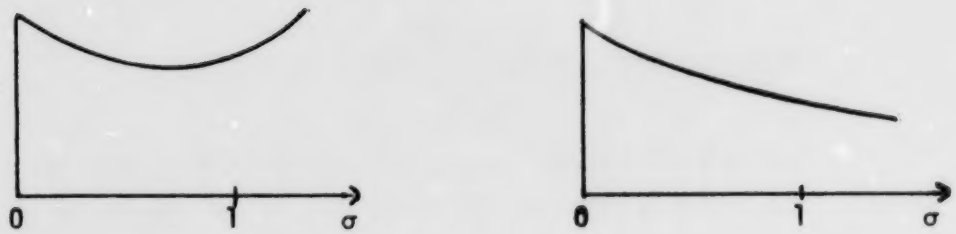
Trouvons maintenant le minimum du membre de droite de (3.58). Sa dérivé est donnée par $- E_k + \bar{n}\, Q\, \sigma\, E_k^2$ .

Lorsque $\sigma = 0$ l'expression est égale à $- E_k < 0$ .

La fonction est donc décroissante à $\sigma = 0$ et son graphe se présente sous l'une des deux formes suivantes:

FIGURE 15

VALEUR DE L'ERREUR EN FONCTION DU PAS




S'il existe $\sigma^* \leq 1$ tel que la dérivé s'annule

$$\sigma^* = \frac{1}{\bar{n}\, Q\, E_k} \leq 1 .$$

Donc si $E_k \geq \frac{1}{\bar{n}\, Q}$ alors $\sigma^* = \frac{1}{\bar{n}\, Q\, E_k}$ et (3.58) devient

$$E_{k+1} \leq \left( 1 - \frac{1}{\bar{n}QE_k} \right) E_k + \tfrac{1}{2}\, \bar{n}\, Q\, \frac{1}{\bar{n}^2 Q^2 E_k^2}\, E_k^2 = E_k - \frac{1}{2\bar{n}Q} .$$

Il s'ensuit que tant que $E_k \geq \frac{1}{\bar{n}Q}$, la valeur de l'erreur décroît d'une quantité au moins égale à $1/(2\bar{n}Q)$. Après un nombre fini d'itérations, nous atteignons donc $E_k < \frac{1}{\bar{n}Q}$ . Nous sommes alors dans le cas où $\sigma^* = 1$. Dans ce cas (3.58) devient

$$E_{k+1} \leq \frac{1}{2}\, \bar{n}\, Q\, E_k^2 .$$

Ceci indique que l'ordre de la convergence est deux.

Pour apprécier la vitesse de convergence dans ce cas, remarquons que $\pi\, Q\, E_{k+1} \leq \frac{1}{2}(\pi\, Q\, E_k)^2$ et après $\ell$ de ces itérations $\pi\, Q\, E_{k+\ell} \leq \frac{1}{2^\ell}\,(\pi\, Q\, E_k)^{2\ell}$ avec $(\pi\, Q\, E_k) < 1$ .

### 3.2.7 Adaptation pour traiter la demande variable

Le modèle P4 donne lieu à des sous-problèmes légèrement différents à cause de la présence des relations de demande variable dans les équations définissant ce modèle (2.35). Le sous-problème du marché r présente la structure suivante

Il existe $u_r$ tel que

$$C_q(h_{Q_r}, h_{\bar{Q}_r}) \left\{ \begin{array}{lll} = u_r & \text{si} & h_q > 0 \\ \geq u_r & \text{si} & h_q = 0 \end{array} \right\} \quad (\forall q \in Q_r) .$$

$$\phi\,(v^r) \qquad \left\{ \begin{array}{lll} = u_r & \text{si} & h_q > 0 . \\ \leq u_r & \text{si} & h_q = 0 . \end{array} \right.$$

$$\sum_{q \in Q_r} h_q - v^r = 0 .$$

$$h_q \geq 0 \quad (\forall q \in Q_r) , \; v^r \geq 0 . \qquad (3.59)$$

Les variables sont dénotées h au lieu de k de façon à conserver l'homogénéité avec les sections précédentes. Ici aussi $C_q(h_{Q_r}, h_{\bar{Q}_r})$ ne dépend que de $h_q$ et sera noté $C_q(h_q)$ et même $C_q$ si le contexte permet d'éliminer $h_q$ sans confusion.

Les sous-problèmes du modèle P4 tant comme ceux du modèle P3, correspondent à des problèmes d'optimisation. Dans ce cas, l'objectif doit être défini de la façon suivante:

$$f(h_{Q_r}, h_{\bar{Q}_r}) = \sum_{q \in Q_r} \int_0^{h_q} C_q(x)\, dx - \int_0^{v^r} \phi(x)\, dx$$

Voyons maintenant les adaptations qu'il faut apporter aux algorithmes discutés précédemment pour les rendre applicables au problème (3.58). Nous nous contentons ici d'énoncer les modifications sans recommencer tous les développements dans ce contexte très voisin du précédent.

Les algorithmes du simplexe convexe et du gradiant réduit sont des méthodes générales, ils peuvent donc traiter cette variante du problème sans difficulté. La base est aussi une matrice $1 \times 1$ pour cette variante. Toutefois, la base peut prendre comme valeur, en plus de 1, la valeur -1 dans le cas où $v^r$ est variable de base.

Pour adapter la méthode du gradiant au modèle P4, il faut remplacer le problème (3.35) définissant la direction de descente par le problème suivant:

$$\text{Min} \sum_{q \in Q_r} C_q d_q - \phi_r \Delta_r$$

$$\sum_{q \in Q_r} d_q - \Delta_r = 0$$

$$\sum_{q \in Q_r} d_q^2 + \Delta_r^2 \leq 1$$

$$d_q \geq 0 \quad (\forall q \in J_k)$$

$$\Delta_r \geq 0 \quad \text{si} \quad v^r = 0 \qquad (3.60)$$

$\Delta_r$ est la modification de la demande.

En étudiant les conditions d'optimalité de Kuhn-Tucker du problème (3.60), on peut montrer que sa solution se calcule de la façon suivante

$$u_j = \frac{\sum_{q \notin J_k} C_q - \phi_r + \sum_{i=1}^{j} C_{q_i}}{\|Q_r\| - \|J_k\| - 1 + j} \quad .$$

Pour calculer $\bar{C}$, on évalue séquentiellement les $u_j$ jusqu'à ce que $C_{q_j} \leq u_j < C_{q_{j+1}}$ et on prend $\bar{C}$ égale au $u_j$ satisfaisant cette relation. Finalement la direction s'obtient ainsi

$$d_q = \begin{cases} \bar{C} - C_q & \text{si} \quad q \in P_k \\ 0 & \text{sinon} \end{cases}$$

$$\Delta_r = \phi_r - \bar{C}$$

Pour adapter la méthode du second ordre au modèle P4, il faut remplacer le problème (3.44) définissant la direction par le problème suivant

$$\text{Min} \sum_{q \in Q_r} (\tfrac{1}{2} S_q d_q^2 + C_q d_q) - \tfrac{1}{2} \phi_r' \Delta_r^2 - \phi_r \Delta_r$$

$$\sum_{q \in Q_r} d_q - \Delta_r = 0$$

$$d_q \geq 0 \quad (\forall q \in J_k)$$

$$\Delta_r \geq 0 \quad \text{si} \quad v^r = 0 \, . \qquad (3.61)$$

En étudiant les conditions d'optimalité de Kuhn-Tucker du problème (3.61),on peut montrer que sa solution se calcule de la façon suivante

$$u_j = \frac{\sum_{q \notin J_k} \frac{C_q}{S_q} - \frac{\phi_r}{\phi_r'} + \sum_{i=1}^{j} \frac{C_{q_i}}{S_{q_i}}}{\sum_{q \notin J_k} \frac{1}{S_q} - \frac{1}{\phi_r'} + \sum_{i=1}^{j} \frac{1}{S_{q_i}}} \quad .$$

$\bar{C}$ est égale à $u_j$ tel que $C_{q_j} \leq u_j < C_{q_{j+1}}$ et la direction devient

$$d_q = \begin{cases} \dfrac{\bar{C} - C_q}{S_q} & \text{si } q \in P_k \, . \\ 0 & \text{sinon} \, . \end{cases}$$

$$\Delta_r = \frac{\phi_r - \bar{C}}{\phi_r'}$$

### 3.2.8 Comparaison numérique des différents algorithmes

Comparons les temps de calcul nécessaires aux quatre algorithmes pour résoudre un sous-problème de choix de routes par les passagers d'un marché. Nous utilisons un sous-problème ayant trois routes dont l'attraction et la disponibilité prennent les valeurs d'un marché type. D'autres études portant sur des problèmes de plus de trois routes et sur d'autres situations d'attraction et de disponibilité ont permis d'observer le même ordonnancement des temps de résolution avec les quatre algorithmes.

La figure suivante présente les performances des quatre algorithmes pour le problème type. Pour cette comparaison,la résolution débute avec la solution réalisable où toutes les variables sont nulles sauf la première. Les quatre algorithmes ont été itérés tant que les nombres de passagers par route n'ont pas été déterminés avec une erreur inférieure à un passager. La diminution du nombre d'itérations et du temps dans l'ordre des algorithmes est assez évidente. Le temps pour le simplexe n'a pas été inclus,car il n'était pas comparable, le programme utilisé n'ayant pas été adapté pour ce type de problème. Le nombre d'itérations par calcul de pas est relativement plus élevé pour les premiers algorithmes,car comme les directions présentent des taux de descente plus faibles, le pas doit être calculé plus précisément pour assurer que la fonction économique décroît à chaque itération et que l'algorithme ne cycle pas. La dernière ligne présente la performance moyenne de la méthode du second ordre pour résoudre tous les sous-problèmes nécessaires à la résolution du problème complet. Le nombre d'itérations est une moyenne pour tous les sous-problèmes ayant nécessité au moins une itération. La performance moyenne est meilleure que l'essai type, car plusieurs des résolutions sont des réoptimisations débutant avec des solutions presque optimales. Le temps moyen de .012 seconde comprend le temps de parcours des sous-problèmes déjà optimal et le temps d'accès à l'information dans un grand réseau.

Tableau 4: Comparaison des temps de calcul

| Algorithme | Nombre d'itération | Nombre d'itération pour le calcul du pas | Temps (sec.) C D C 173 |
|---|---|---|---|
| Simplexe-convexe | 19 | - | - |
| Gradiant réduit | 11 | 155 | .4 |
| Gradiant | 7 | 40 | .1 |
| Second ordre | 2 | 6 | .02 |
| Second ordre (Implantation) | 1.1 | 3 | .012 |

CHAPITRE 4

## EXPERIMENTATION DES MODELES DE REPARTITION DES PASSAGERS

Plutôt que d'expérimenter tous les modèles avec un problème hypothétique ou réel de dimension réduite, nous avons préféré commencer l'expérimentation sur le réseau d'une grande compagnie aérienne en vue de démontrer clairement le potentiel d'utilisation. L'expérimentation a jusqu'ici été limitée à celle du premier modèle à cause de l'ampleur de l'entreprise et de la non disponibilité des coefficients d'attraction sous la forme utilisée par les trois derniers modèles. Un modèle de répartition des passagers en fonction de l'attraction des routes devra être calibré pour permettre l'utilisation des trois derniers modèles.

Le premier modèle développé qu'on a nommé MAPUM* donne d'excellents résultats. Les trois autres modèles lui sont théoriquement supérieurs, ils devraient donc donner des résultats encore meilleurs. Cependant, seul les essais futurs permettront de quantifier les améliorations qu'apportent les derniers modèles.

### 4.1 Plan d'expérience

Comme problème-test nous avons utilisé l'ensemble du réseau d'Air Canada. La demande et l'horaire de l'hiver 76 ont été utilisés. Les résultats du modèle peuvent ainsi être comparés avec l'utilisation des avions observée durant cette période. Les résultats sont aussi comparés avec la répartition des passagers obtenue par Air Canada avec leur version du modèle ASPEM. Ce dernier modèle "Airline Schedule Planning and Evaluation Model" a été développé par Douglas Aircraft et est utilisé par plusieurs compagnies aériennes.

---

*MAPUM : Modèle Allocation de Passagers Université de Montréal

La deuxième étape de l'évaluation consiste à étudier la répartition dans le réseau actuel de la demande prévue pour 82-83. Cette répartition est faite en parallèle par les modèles MAPUM et ASPEM. Nous tenterons de dégager l'utilité des résultats des 2 modèles pour le réajustement du service offert en fonction de la demande prévue.

### 4.1.1 Description du réseau utilisé

Air Canada opère à la fois un réseau international s'étendant aux Etats-Unis, aux Antilles et en Europe et un vaste réseau intérieur comprenant une trentaine de villes fortement connectées entre elles. Le cheminement des passagers dans ce réseau est assez complexe car deux villes sont généralement reliées par plusieurs itinéraires dans l'espace et le temps.

La Figure 16 présente des cartes simplifiées des réseaux international et nord-américain; tous les vols reliant deux villes sont représentés par un seul arc . Le réseau complet comprend 76 villes et 629 paires origines-destinations. Ces origines-destinations explicites comprennent toutes les paires de villes entre lesquelles voyagent au moins un passager par jour. Les passagers des paires de villes moins importantes sont remplacés par des passagers supplémentaires pour les paires origines destinations intermédiaires de leur itinéraire probable.

L'horaire considéré comprend 329 vols dont 202 sont des vols directs, 73 sont des vols d'une escale et 54 sont des vols comportant de 2 à 5 escales. Ces vols comprennent un total de 539 arcs. Ils offrent aux différentes origines-destinations un total de 800 chaînes sans correspondance et 700 chaînes avec correspondance en retenant toutes les chaînes avec correspondance transportant au moins un passager par jour.

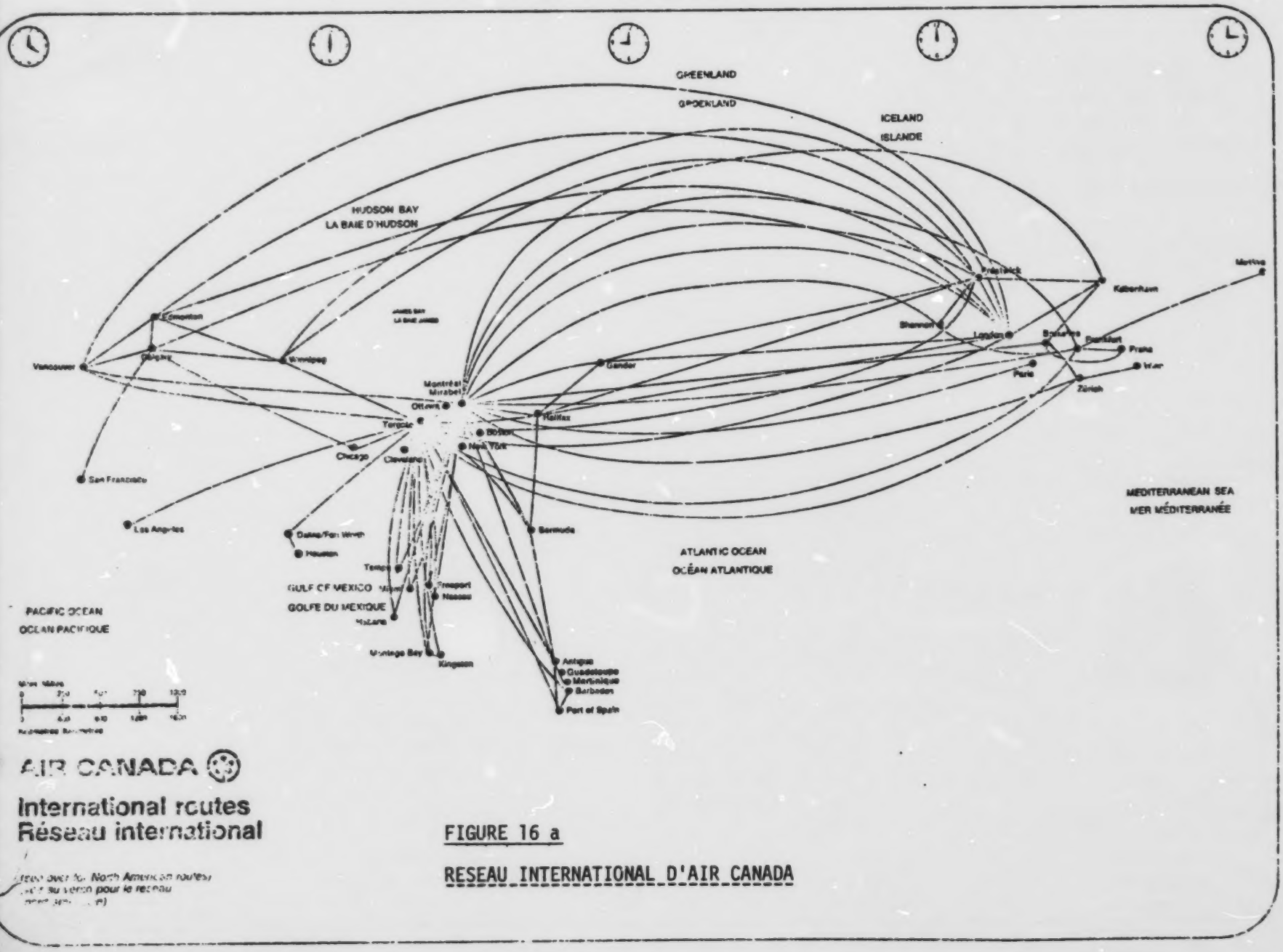



FIGURE 16 a

RESEAU INTERNATIONAL D'AIR CANADA

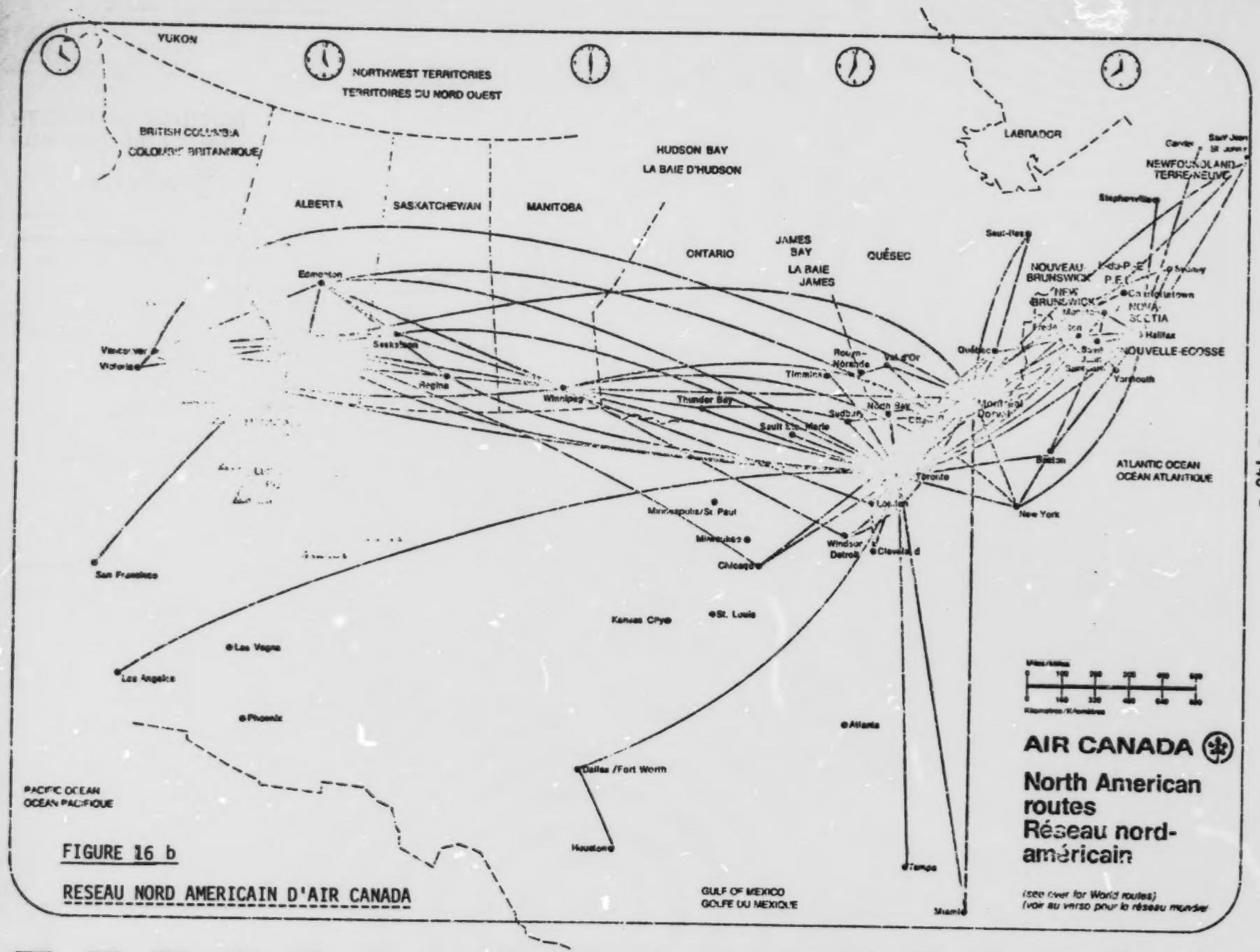



FIGURE 16 b

RESEAU NORD AMERICAIN D'AIR CANADA

### 4.1.2 Calibration du modèle

L'objectif du modèle P1, tel que défini à la section 2.4, comporte un terme mesurant les revenus perdus à cause des passagers excédentaires et un terme mesurant le coût d'utiliser des routes moins attrayantes. Ces coûts ont été déterminés en se préoccupant surtout de l'ajustement des coûts entre les marchés. Par contre, la constante K pondérant l'effet relatif de l'attraction et de la disponibilité n'a pas été précisée. Nous déterminons ici ce paramètre de façon à reproduire l'effet relatif de l'attraction et de la disponibilité ayant déterminé l'utilisation des avions à l'hiver 76.

En augmentant le paramètre K, nous augmentons le coût relatif des routes moins attrayantes. La solution associée comprend donc plus de passagers sur les routes plus attrayantes, car il devient relativement moins coûteux de risquer de les refuser faute de place que de les détourner sur une route moins attrayante. Ainsi, les arcs les plus utilisés le deviennent encore plus avec une augmentation de K tandis que les arcs moins utilisés le deviennent encore moins.

Pour ajuster le paramètre K, nous procédons donc ainsi; à partir de l'utilisation des avions observés, nous partageons les arcs en 3 sous-ensembles; ceux très occupés, ceux moyennement occupés et ceux peu occupés. Ces sous-ensembles comprennent respectivement les arcs pour lesquels $(c_a - u_a)/\sigma_a \in (-\infty,1), (1,2), (2,\infty)$ où $c_a$ est la capacité de l'avion, $u_a$ le nombre moyen de passagers observé et $\sigma_a$ son écart-type. Pour chacun de ces ensembles, nous considérons la différence moyenne entre le nombre de passagers affecté par le modèle pour une valeur de K et le nombre de passagers observé. Pour l'ensemble des arcs moyennement occupés, la différence moyenne reste pratiquement nulle pour les valeurs de K considérées. Un k trop grand rend cette différence positive pour l'ensemble des arcs très occupés et négative pour les arcs peu occupés.

Un K trop petit produit l'effet inverse. La valeur de K est donc choisie de façon à rendre presque nulles les différences pour les 3 sous-ensembles d'arcs. Le tableau 5 présente ces différences pour différentes valeurs de K. Remarquons de plus que la valeur de K = 1/3 choisie par cette méthode est une de celle qui minimise l'erreur globale.

Tableau 5 Effets du paramètre K

| K | Erreur moyenne pour chaque ensemble d'arcs | | | Moyenne pour tous les arcs de la valeur absolue de l'erreur |
|---|---|---|---|---|
| | Peu occupés | Moyennement occupés | Très occupés | |
| 1 | -3.8 | 0 | 4.1 | 11.6 |
| $\frac{1}{2}$ | -1.8 | .3 | 1.1 | 11.3 |
| $\frac{1}{3}$ | - .5 | .3 | - .7 | 11.3 |
| $\frac{1}{4}$ | .4 | .1 | -1.4 | 11.3 |
| $\frac{1}{10}$ | 1.7 | .4 | -4.3 | 12.0 |

## 4.2 Validation du modèle avec la demande d'une saison passée

Nous avons utilisé comme demande les passagers qui ont voyagé au cour de l'hiver 76. Le remplissage des avions calculé par le modèle est ensuite comparé avec la situation observée. Cette expérience permet d'étudier l'erreur d'ajustement du modèle sans la présence d'erreur sur la prévision de la demande.

### 4.2.1 Résultats globaux

La figure 17 présente pour chaque arc le résultat de MAPUM en fonction du nombre de passagers observé . La concentration des points près de la diagonale illustre la qualité des prédictions du modèle. Ici, la concentration est un peu masquée par la superposition des points. Ce type de graphique sera repris plus loin avec seulement un sous-ensemble des arcs, ce qui permettra de mieux visualiser la relation.

Une seconde mesure de la qualité de la prédiction de MAPUM est donnée par la figure 18. Cette figure présente la distribution de la différence entre l'estimé de MAPUM et le nombre de passagers observé pour chaque arc. La moyenne des différences en valeur absolue est de 11.3, et la différence moyenne .3 .

La figure suivante présente avec la même mesure, la qualité de la prédiction du modèle ASPEM. Avec ce modèle, la moyenne des différences en valeur absolue est 23.4, ce qui est plus du double de l'erreur moyenne de MAPUM. De plus, en comparant les distributions des figures 18 et 19, nous constatons que le pourcentage des arcs où la prédiction est exacte à ± 5 passagers est de 36% pour MAPUM contre 18% seulement pour ASPEM. D'autre part, l'erreur est supérieure à ± 15 passagers pour 50% des arcs avec ASPEM contre seulement 25% avec MAPUM.

FIGURE 17: NOMBRE DE PASSAGERS ESTIME PAR MAPUM EN FONCTION DU NOMBRE DE PASSAGERS OBSERVE POUR CHAQUE ARC.

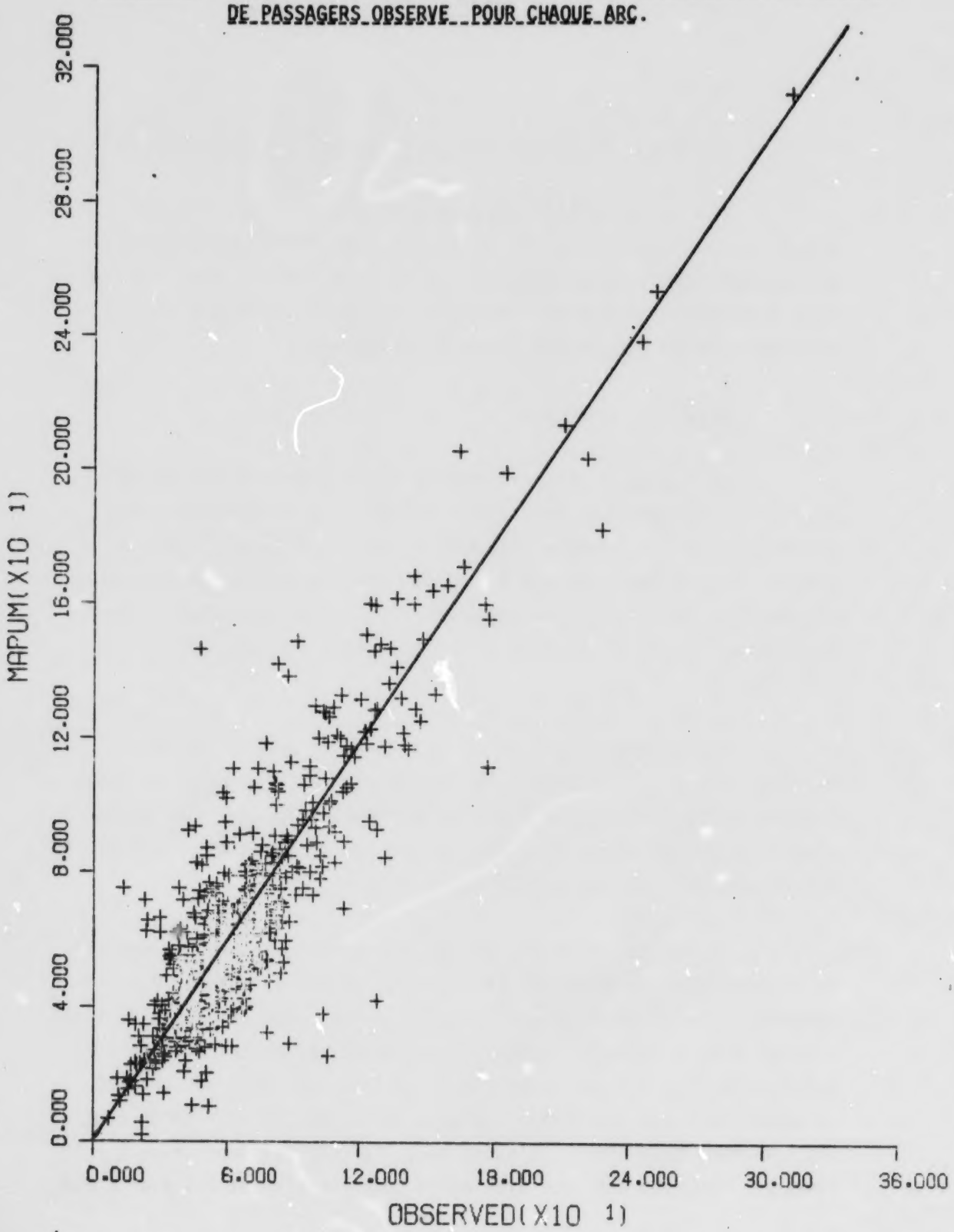


FIGURE 18: DISTRIBUTION DE LA DIFFERENCE ENTRE L'ESTIME DE MAPUM ET LE NOMBRE DE PASSAGERS OBSERVE POUR CHAQUE ARCS

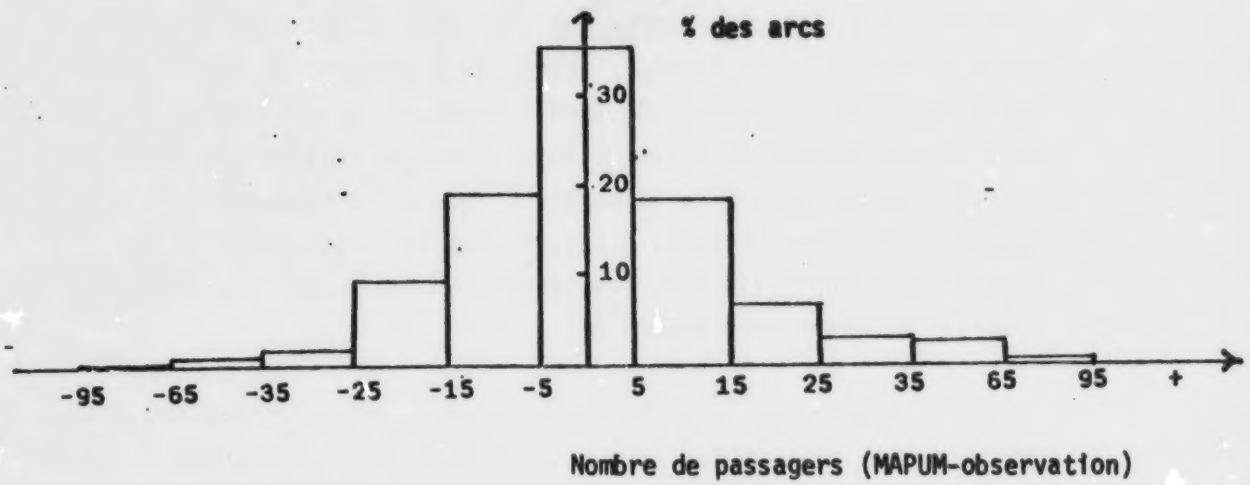
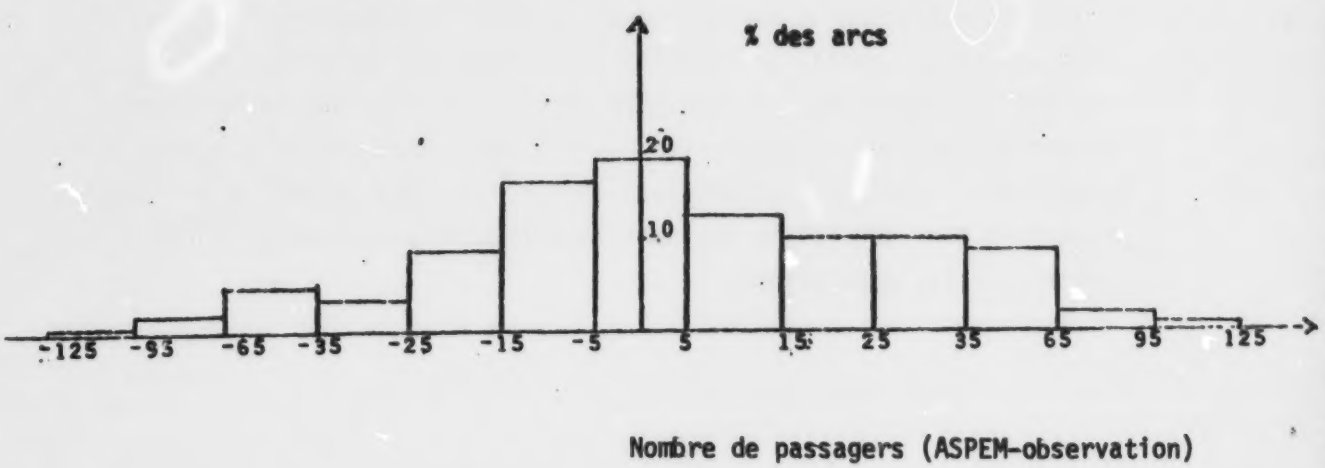



FIGURE 19: DISTRIBUTION DE LA DIFFERENCE ENTRE L'ESTIME D'ASPEM ET LE NOMBRE DE PASSAGERS OBSERVE POUR CHAQUE ARC.



La table 6 présente la variation de l'erreur de MAPUM en fonction des caractéristiques des arcs. D'abord le modèle est beaucoup plus précis pour les arcs ayant un coefficient d'utilisation élevé. Pour ces arcs, la répartition des passagers est fortement dépendante de la disponibilité, tandis que pour les arcs moins occupés elle dépend surtout de l'attraction. Ainsi, la qualité des résultats illustre la pertinence de la notion de passagers excédentaires pour modéliser la disponibilité. La faiblesse des résultats pour les arcs moins occupés vient de l'utilisation d'un modèle simpliste et pourrait sûrement être améliorée.

L'effet de la distance se résume à une surestimation du nombre de passagers pour les arcs longs. Ceci correspond probablement au déplacement des passagers de routes avec escale et correspondance vers des routes directes. Des études plus complètes seraient nécessaires avant de réajuster la calibration sur ce point étant donné qu'il n'a pas beaucoup d'arcs longs et que l'effet est faible.

Finalement, on retrouve des erreurs systématiques importantes dans l'utilisation de certain type d'avions. Les R727 sont sous-utilisés suivant le modèle et les L1011 le sont trop. L'erreur pour ces avions desservant Montréal-Toronto provient du fait que le modèle ne tient pas compte de la priorité d'utilisation accordée au R727 par la compagnie. Les 747 sont sous-utilisés et les D8SF le sont trop. Ceci pourrait être corrigé en incluant dans le modèle un facteur d'attraction pour les gros avions.

TABLE 6 REPARTITION DE L'ERREUR PAR SOUS-ENSEMBLE D'ARCS

| Sous-Ensembles d'Arcs | Nombre d'arcs | Erreur moyenne | Erreur absolue moyenne |
|---|---|---|---|
| Coefficient de remplissage | | | |
| - Faible | 144 | -.5 | 13.8 |
| - Moyen | 326 | .3 | 11.0 |
| - Elevé | 69 | -.7 | 8.6 |
| Distance (en heure de vol) | | | |
| - Moins de 1.5 heure | 415 | -.3 | 11.2 |
| - Entre 1.5 et 4 heures | 81 | 1.7 | 11.3 |
| - Plus de 4 heures | 43 | 3.3 | 13.4 |
| Type d'avion | | | |
| - DC9 | 76 | .2 | 8.0 |
| - D9SF | 289 | -.6 | 9.3 |
| - A727 | 49 | 1.2 | 11.4 |
| - DC8F | 7 | -1.2 | 3.8 |
| - R727 | 28 | -14.6 | 22.8 |
| - B747 | 3 | -16 | 28.4 |
| - D8SF | 51 | 6.6 | 15.2 |
| - L1011 | 28 | 15.2 | 27.9 |
| - A747 | 8 | -5.4 | 9.2 |

### 4.2.2 Résultats pour le sous-réseau de l'Ouest Canadien

Ce sous-réseau a été étudié, car il intéressait particulièrement la compagnie Air Canada à cette époque. On continue de faire la répartition des passagers dans l'ensemble du réseau même si on discute seulement des résultats obtenus pour le sous-réseau.

Ce sous réseau comprend 7 villes, des liaisons directes entre 31 paires de ces villes et des fréquences des services assez élevées entre chaque paire de villes. La figure 20 présente le réseau avec les nombres de liaisons directes dans le sens ouest-est. Les arcs est-ouest ont également été considérés même s'ils n'apparaissent pas sur cette figure. Dans le triangle YYZ-YOW-YUL, seul les vols ayant une origine ou une destination plus à l'ouest sont considérés. Nous obtenons ainsi un réseau de 150 arcs desservant les marchés de l'Ouest Canadien.

Les figures 21 et 22 présentent respectivement les nombres de passagers estimés par MAPUM et ASPEM en fonction du nombre de passagers observé pour chacun des arcs de l'Ouest Canadien. La qualité des prévisions de MAPUM par rapport à celles de ASPEM est évidente.

### 4.2.3 Résultats détaillés pour certaines origines-destinations

Les figures 23 à 27 présentent des exemples de résultats par paires origines-destinations. Ces figures présentent les passagers locaux par route et les passagers totaux par arc pour MAPUM, ASPEM et les observations.

On peut constater que les résultats de MAPUM varient de façon simulaire aux nombres de passagers observés. Même si on observe des écarts entre les nombres déterminés par MAPUM et ceux observés, on peut remarquer que les arcs que MAPUM prévoit très occupés et ceux qu'il prévoit peu occupés le sont effectivement, ce qui n'est pas le

FIGURE 20: OUEST CANADIEN: NOMBRE DE DEPARTS PAR ARC.

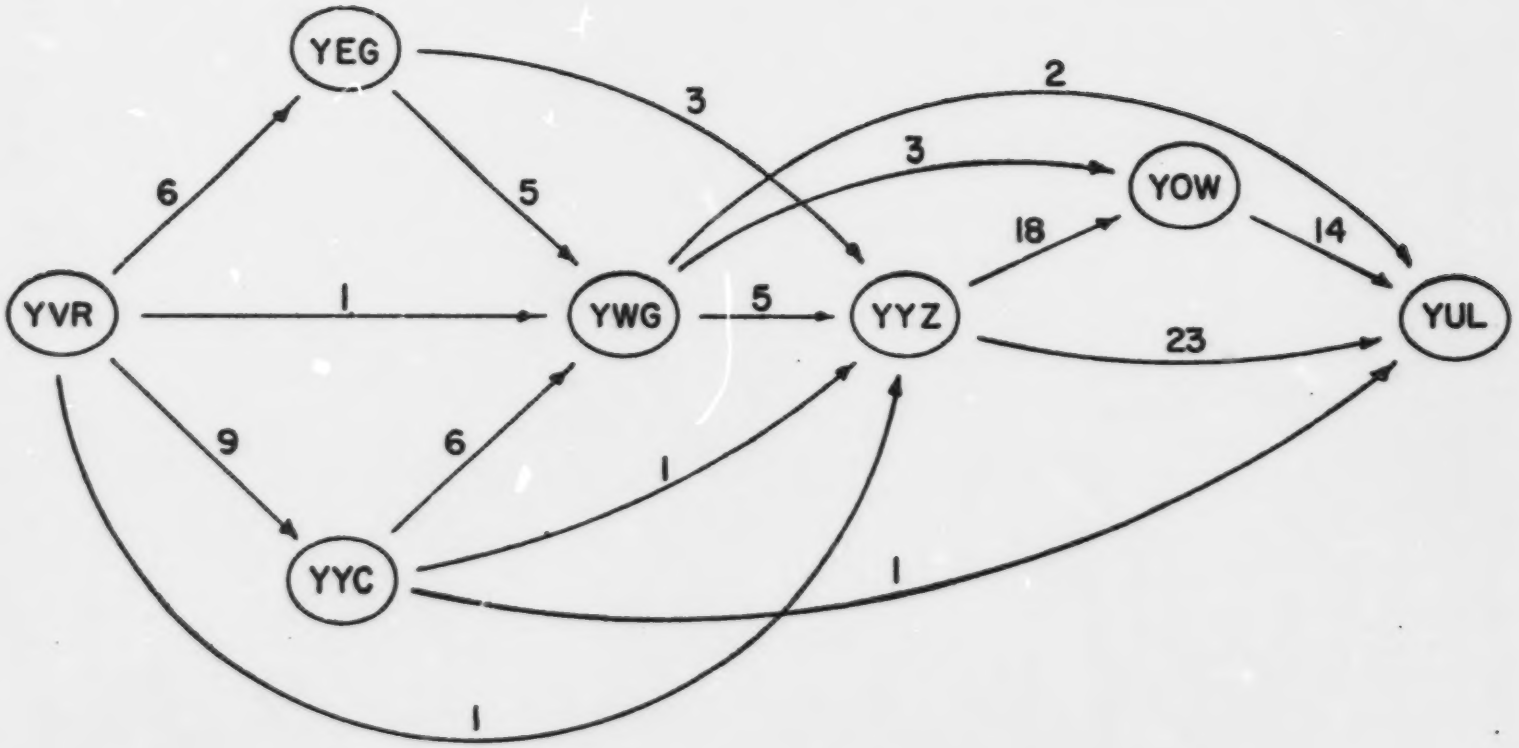


FIGURE 21: NOMBRE DE PASSAGERS ESTIME PAR MAPUM EN FONCTION DU NOMBRE DE PASSAGERS OBSERVE POUR LES ARCS DE L'OUEST CANADIEN.

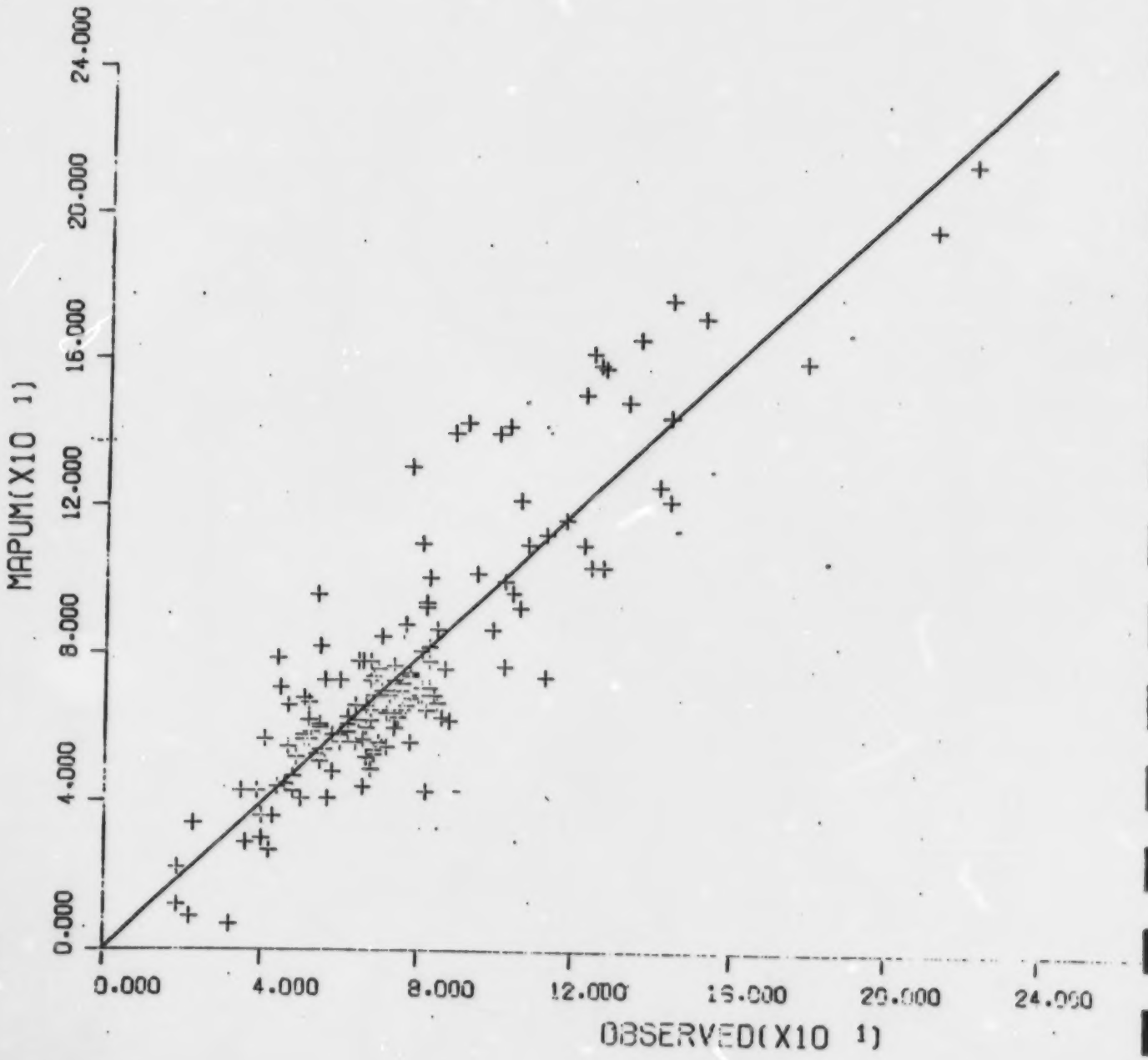
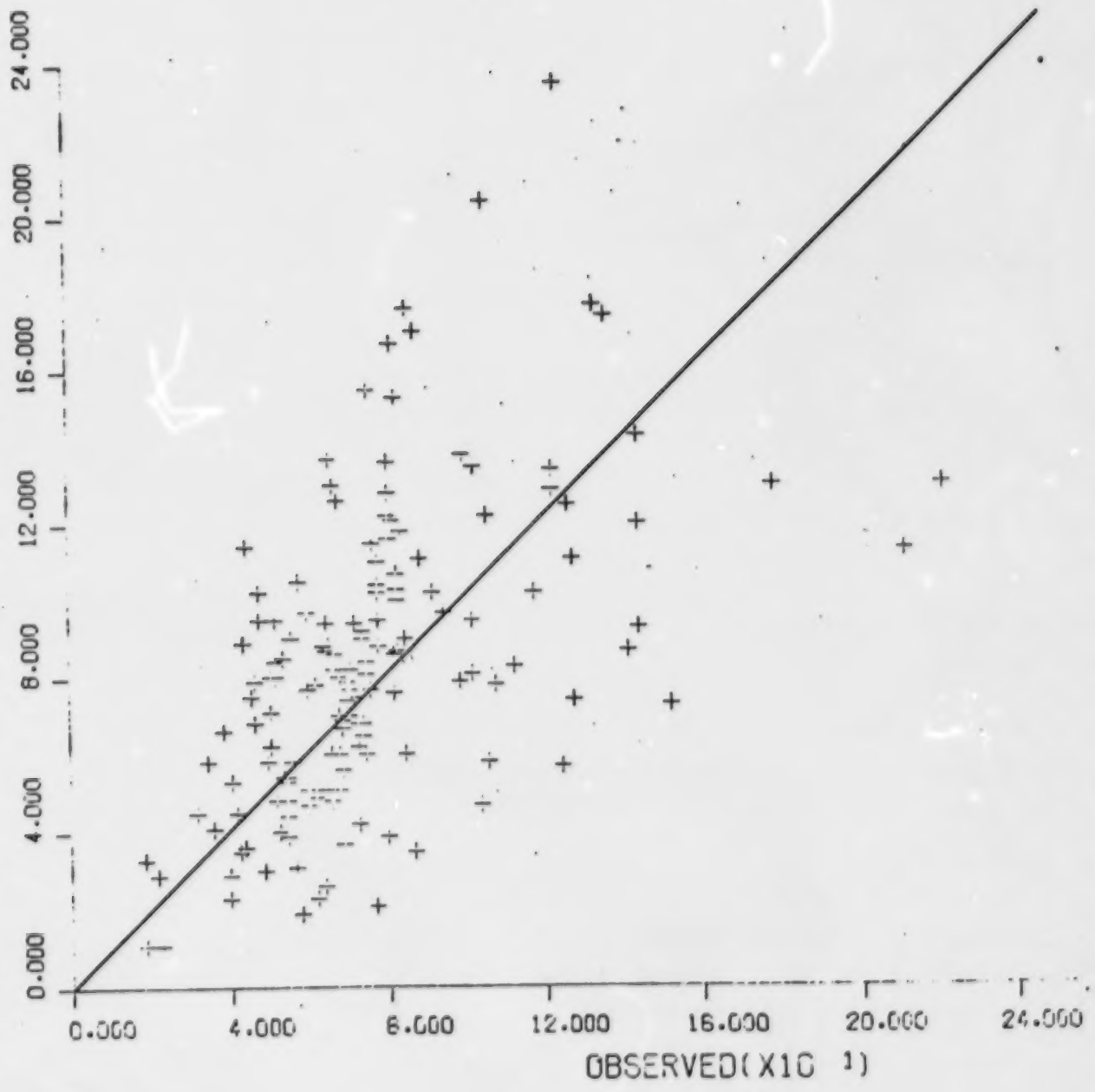


FIGURE 22: NOMBRE DE PASSAGERS ESTIME PAR ASPEM EN FONCTION DU NOMBRE DE PASSAGERS OBSERVE POUR LES ARCS DE L'OUEST CANADIEN.



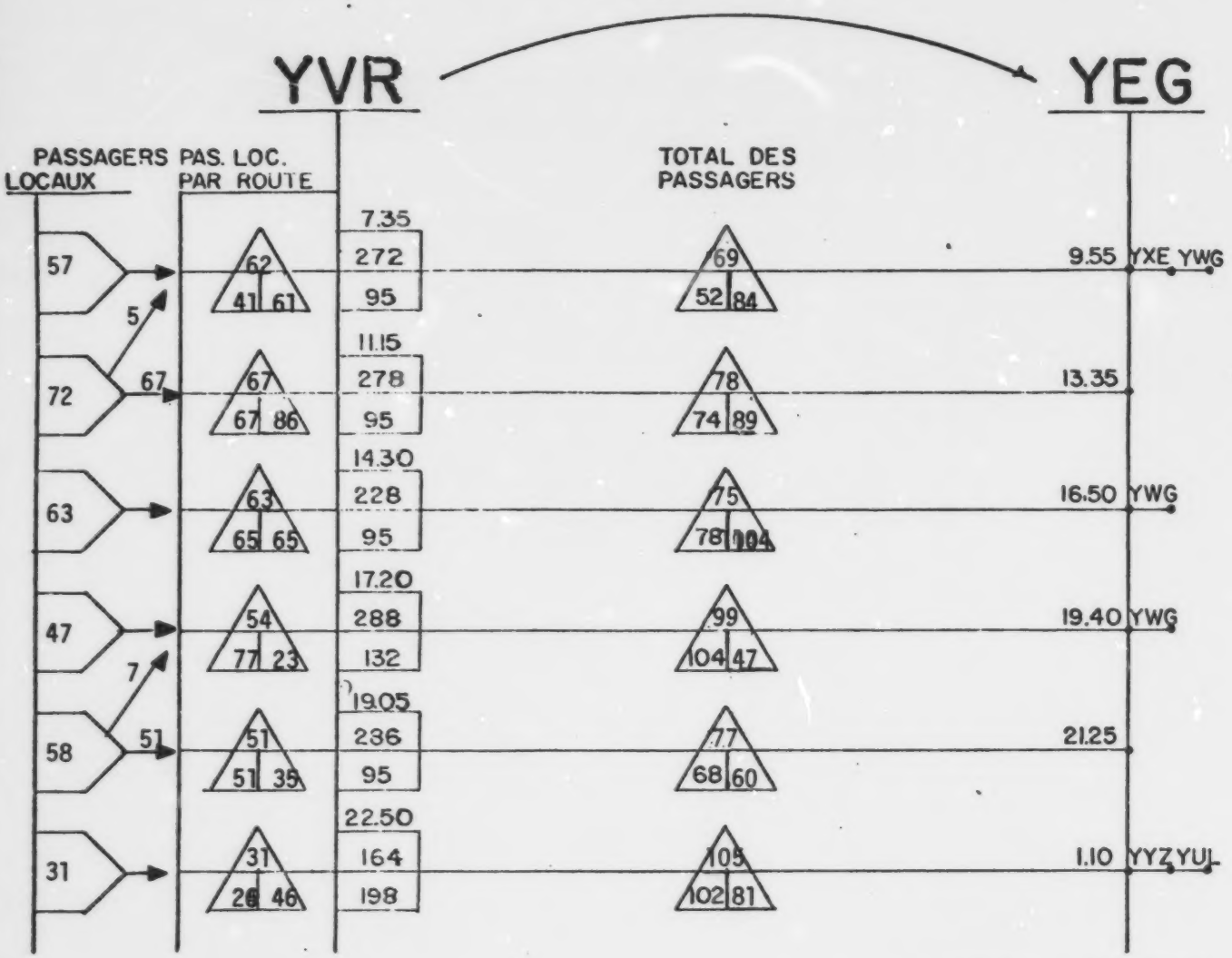



FIGURE 23: RESULTATS DETAILLES DE 76 POUR L'O/D YVR-YEG

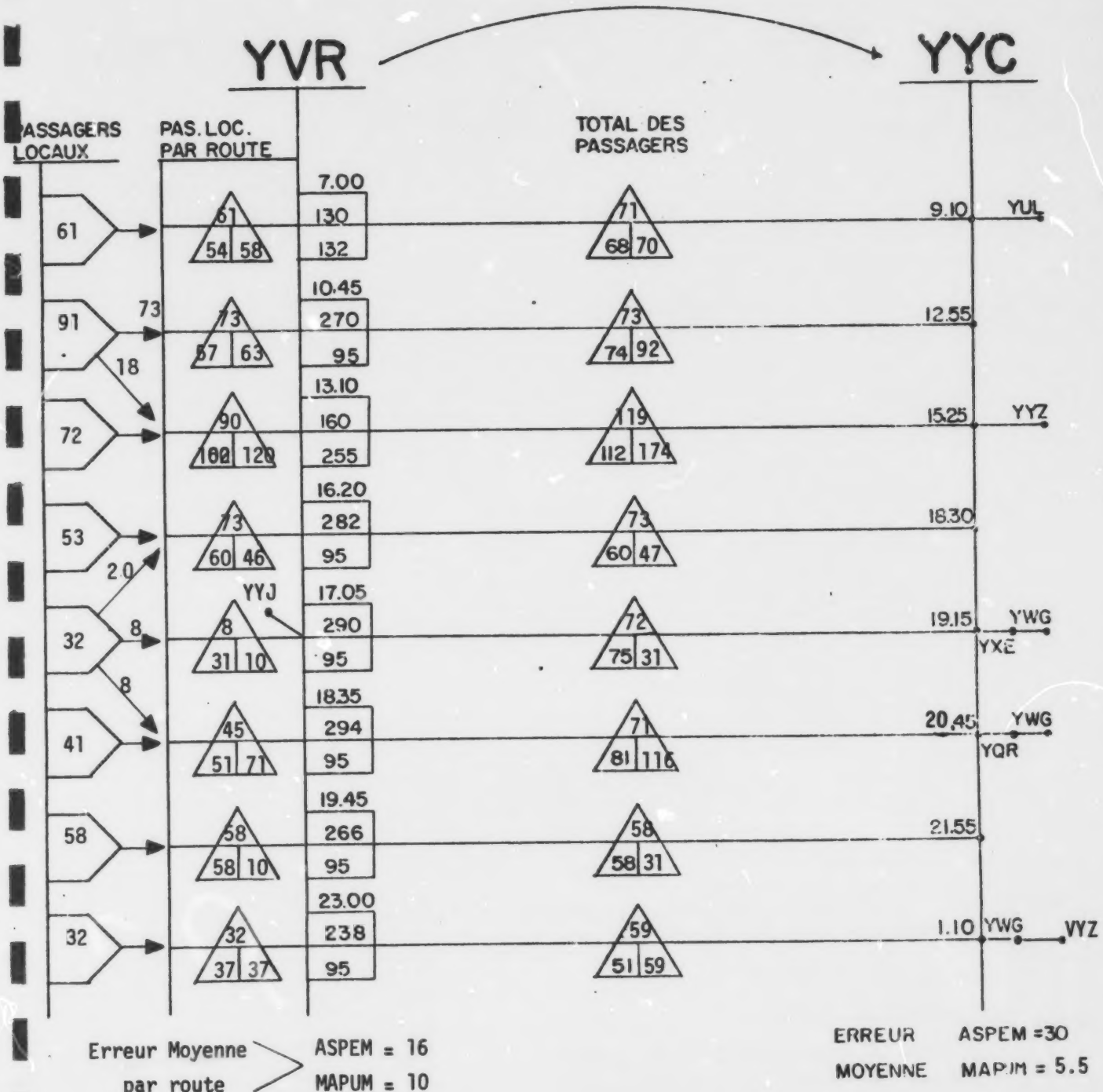
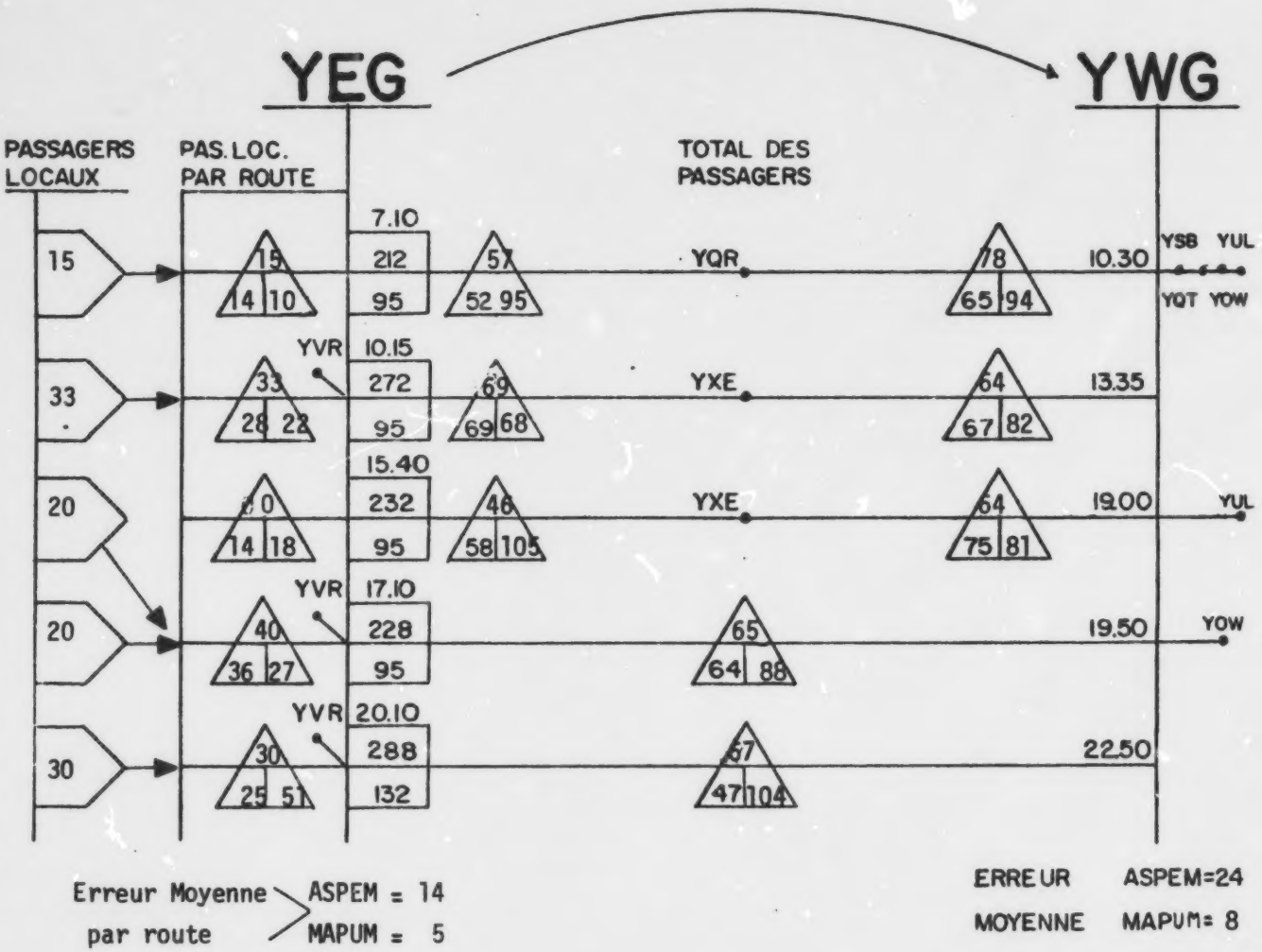



FIGURE 24

RESULTATS DETAILLES DE 76 POUR L'O/D YVR-YYC



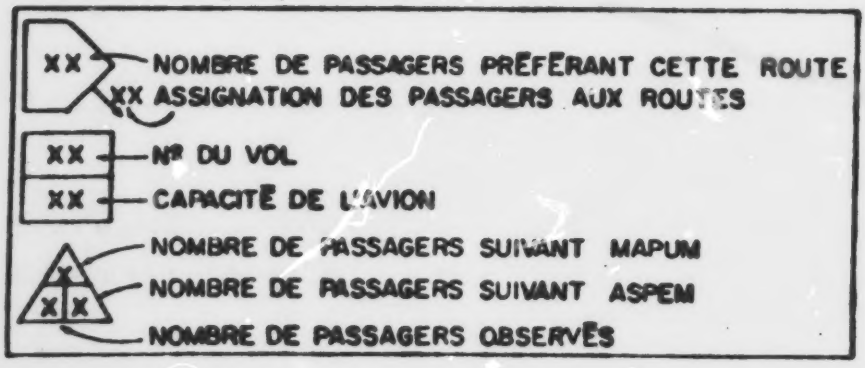




FIGURE 25

RESULTATS DETAILLES DE 76 POUR L'O/D YEG-YWG

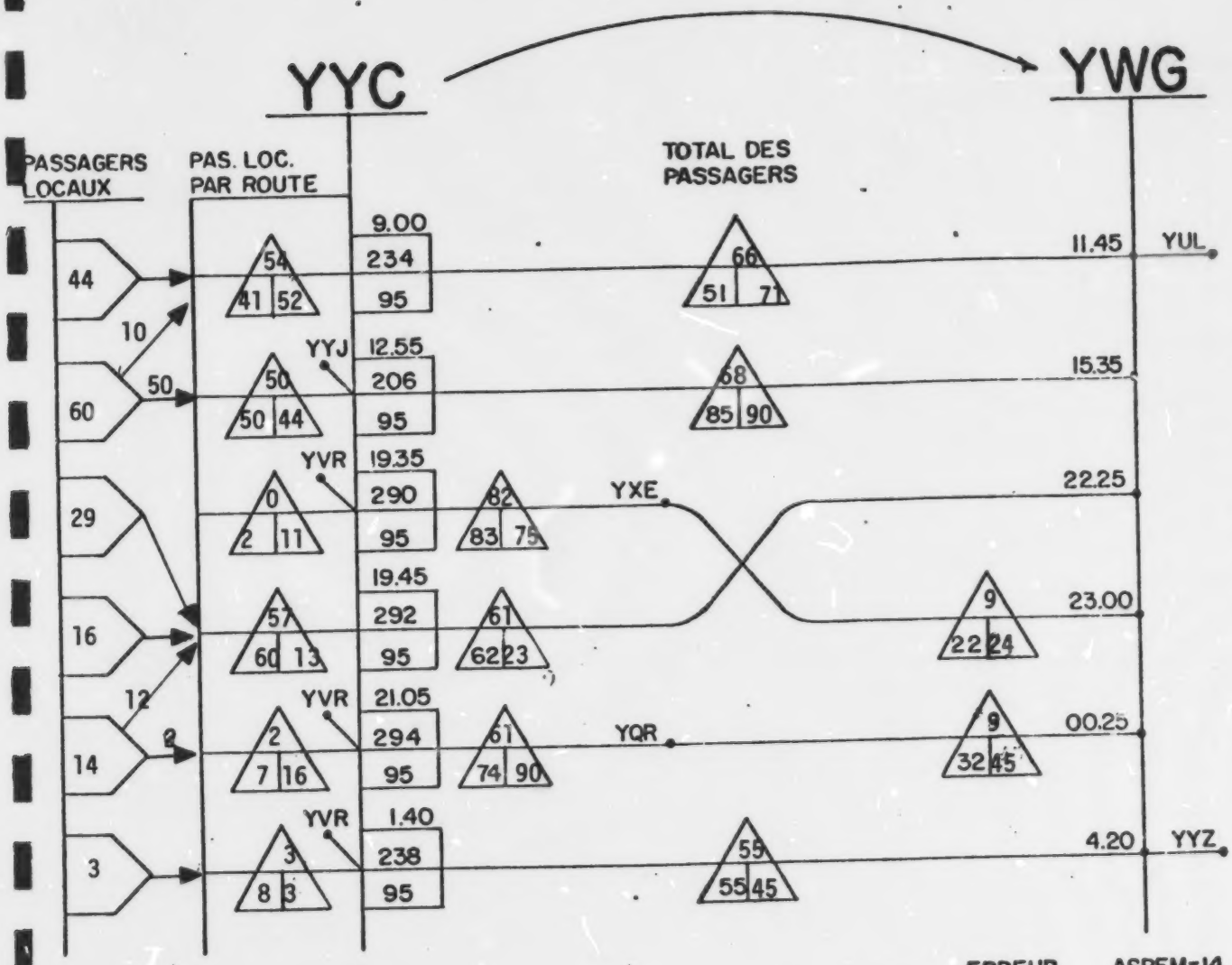



FIGURE 26

RESULTATS DETAILLES DE 76 POUR L'O/D YYC-YWG

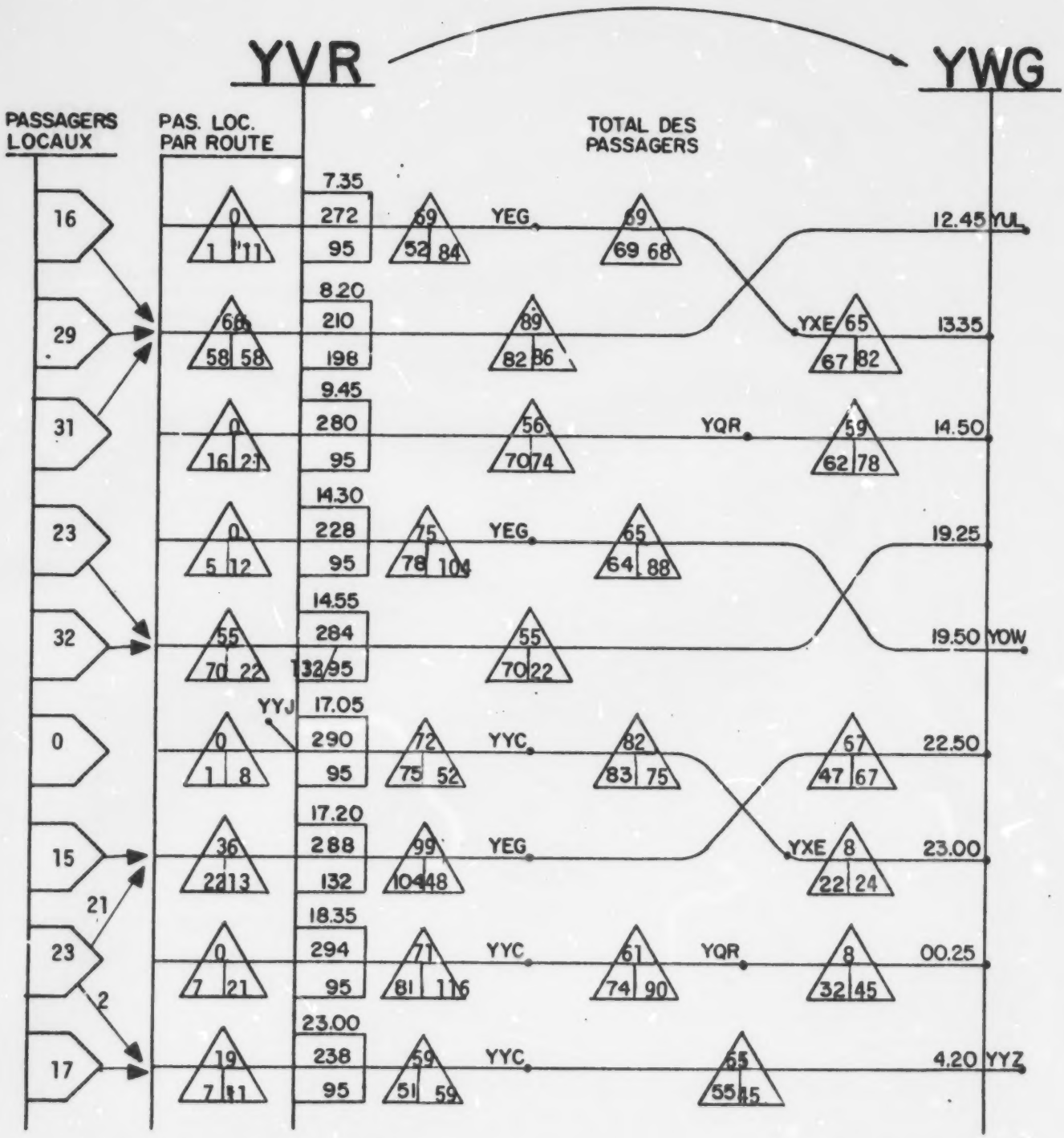



FIGURE 27

RESULTATS DETAILLES DE 76 POUR L'O/D YVR-YWG

cas avec ASPEM. Une partie importante des erreurs vient des distributions des heures de départ préférées. Ces courbes fournies par Air Canada indiquent trop de préférences de départ l'avant-midi et pas assez l'après-midi.

Le tableau 7 compare les préférences de départ l'avant-midi et l'après-midi indiquées par les distributions utilisées et les nombres de départ observés pendant ces périodes.

Tableau 7 Ecart entre les distributions d'heures de départ désirées et les heures de départ observés.

| Origine Destination | Distributions Utilisées | | Départs Observés | |
|---|---|---|---|---|
| | AM | PM | AM | PM |
| YUR-YEG | 129 | 110 | 108 | 142 |
| YUR-YYC | 152 | 157 | 111 | 193 |
| YEG-YWG | 48 | 40 | 42 | 50 |
| YYC-YWG | 104 | 45 | 91 | 62 |
| YVR-YWG | 76 | 70 | 75 | 98 |
| Total | 509 | 422 | 427 | 545 |

## 4.3 Utilisation du modèle avec une estimation de la demande

La demande de 82-83 fut estimée en augmentant la demande de 76 du pourcentage de croissance différent pour chaque origine-destination. Les passagers furent distribués dans le réseau défini par les vols de 76. Les figures 28 à 32 présentent les résultats pour les cinq origines-destinations étudiées précédemment. Les nombres de passagers observés sont ceux de 76, ceux de 82-83 étant évidemment inconnus.

Un premier examen permet de constater que les nombres de passagers par arc estimés pour MAPUM sont supérieurs aux nombres de passagers observés en 76, ce qui n'est pas toujours le cas avec ASPEM.

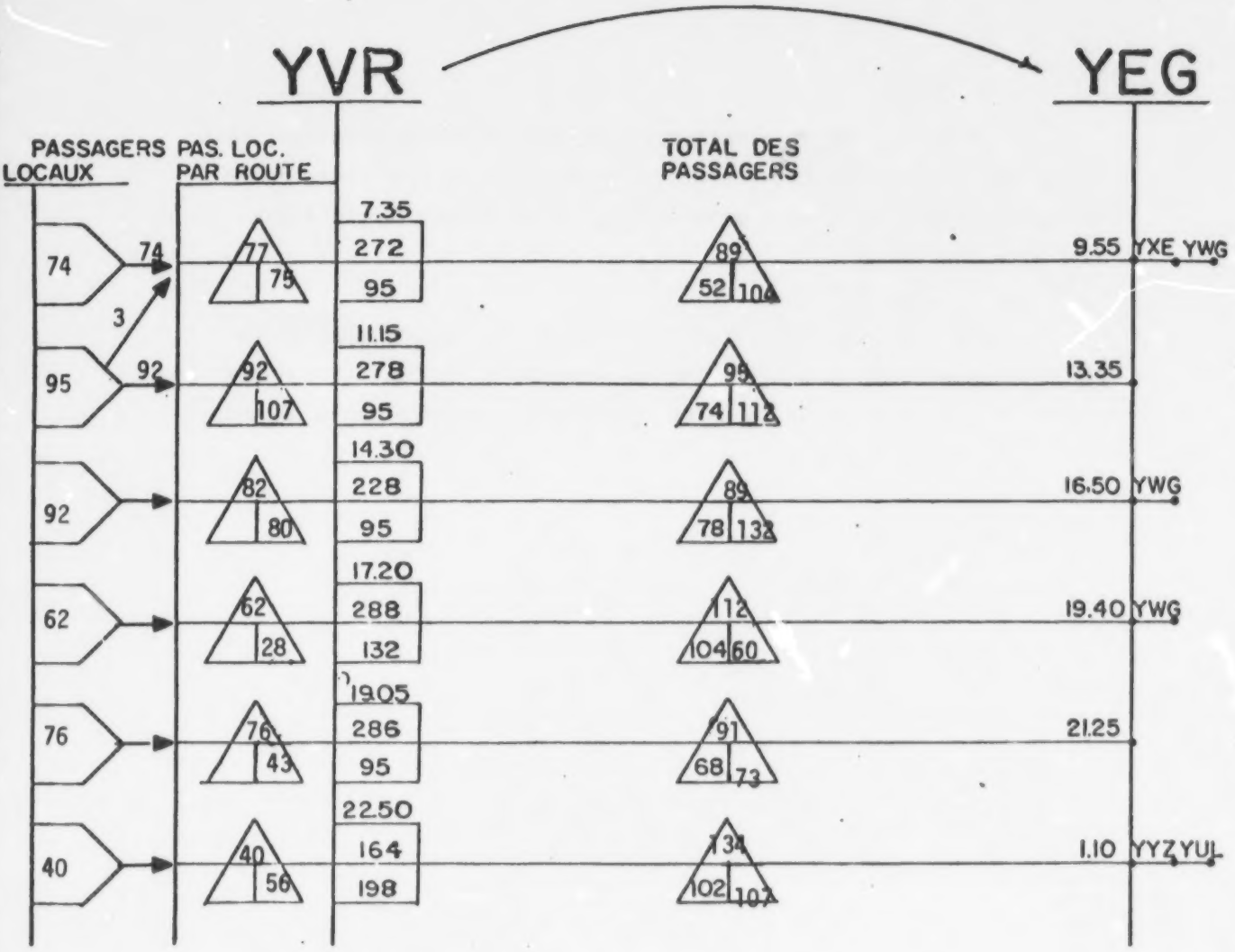
Voyons maintenant l'information que le planificateur obtient au niveau d'un marché.

Pour YVR-YEG les avions sont uniformément très occupés. Il n'y a pas avantage à changer l'heure des vols, par contre en considérant les nombres de passagers refusés,faute de place, le planificateur doit décider s'il y a lieu d'augmenter la capacité offerte.

Pour YVR-YYC nous observons un déplacement des passagers de 17 heures vers les 2 vols précédents et les deux vols suivants. Le remplissage des avions est satisfaisant. Le planificateur devra décider,s'il y a lieu,de déplacer des sièges vers 17 heures pour augmenter la satisfaction des passagers en considérant aussi la concurrence et la disponibilité des avions.

Pour YEG-YWG le remplissage des avions est satisfaisant mais il y a déplacement des passagers de l'avant-midi vers le reste de la journée. Une décision similaire à celle de l'exemple précédent doit être considérée.

Il est aussi intéressant d'observer la ressemblance entre les déplacements des passagers locaux entre 76 et 82-83. Ce qui démontre la stabilité de la solution lors d'une variation de la demande.




XX — NOMBRE DE PASSAGERS PRÉFÉRANT CETTE ROUTE
XX ASSIGNATION DES PASSAGERS AUX ROUTES

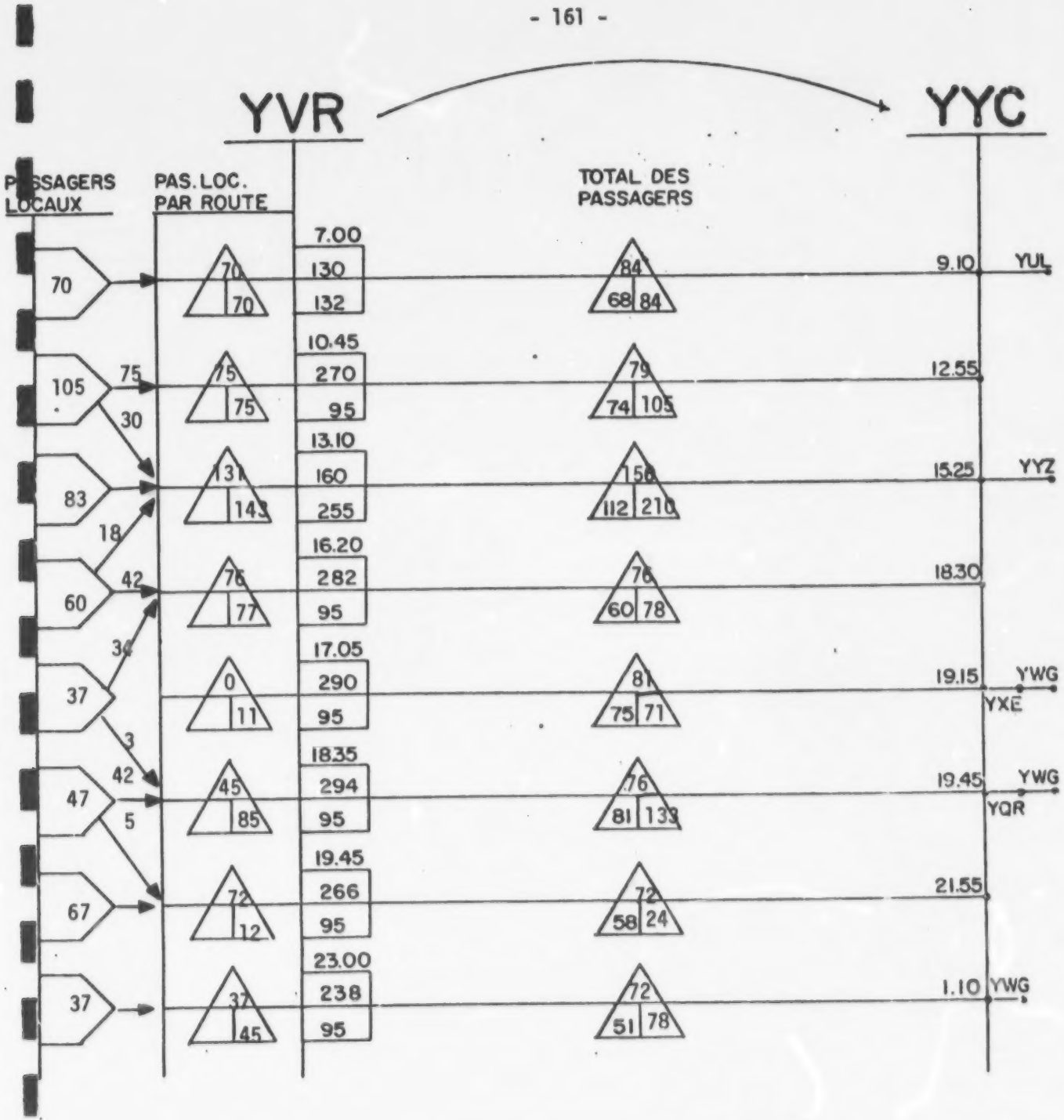
XX — Nº DU VOL
XX — CAPACITÉ DE L'AVION

NOMBRE DE PASSAGERS SUIVANT MAPUM 82 - 83
NOMBRE DE PASSAGERS SUIVANT ASPEM 82 - 83
NOMBRE DE PASSAGERS OBSERVÉS 76

Coefficient d'utilisation 86%

FIGURE 28

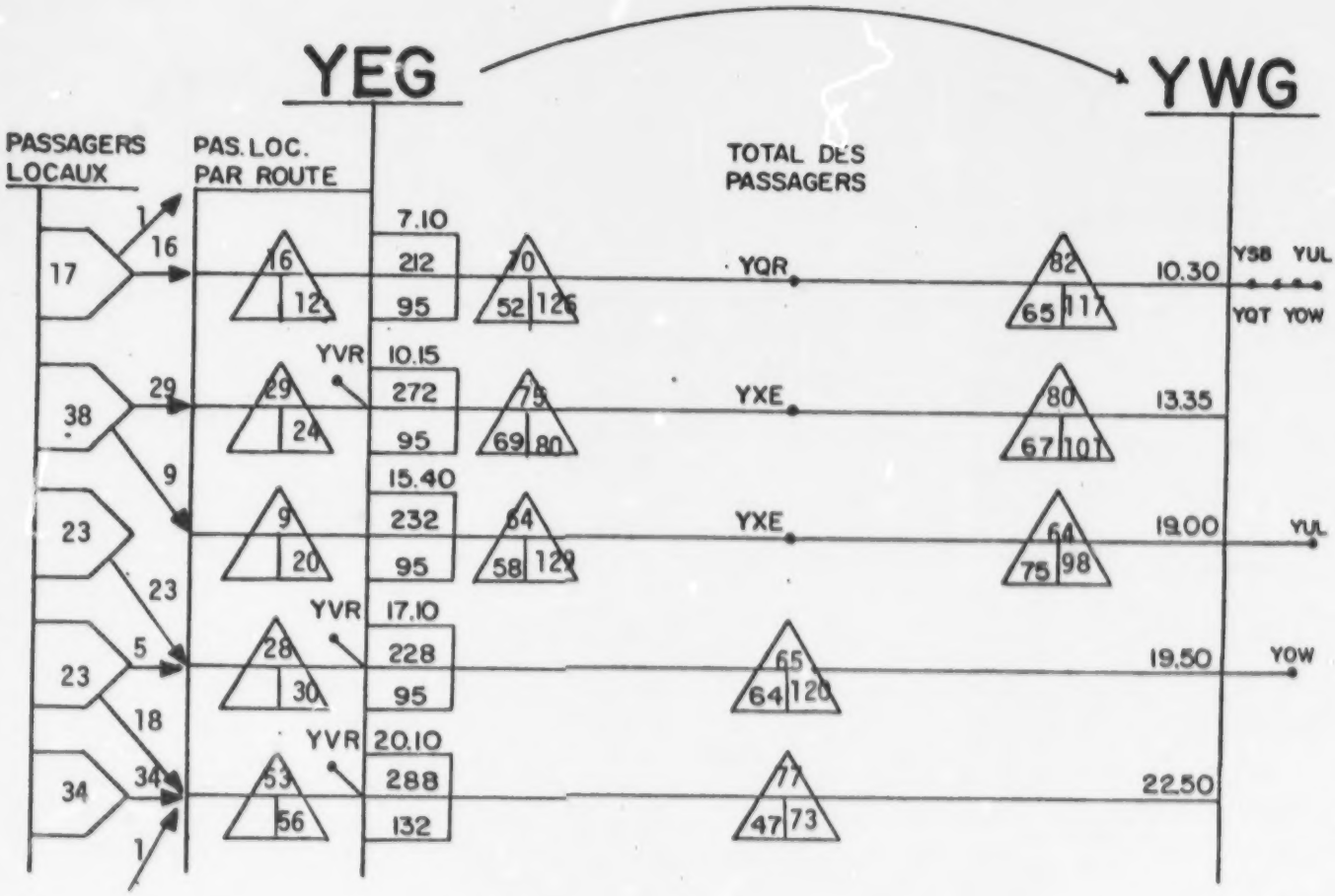
CIRCULATION ESTIMEE POUR 82-83 POUR L'O/D YVR-YEG




Coefficient d'utilisation 73%

FIGURE 29

CIRCULATION ESTIMEE POUR 82-83 POUR L'O/D YVR-YYC




Coefficient d'utilisation 72%

XX — NOMBRE DE PASSAGERS PRÉFÉRANT CETTE ROUTE
XX ASSIGNATION DES PASSAGERS AUX ROUTES

XX — Nº DU VOL
XX — CAPACITÉ DE L'AVION

NOMBRE DE PASSAGERS SUIVANT MAPUM 82 - 83
NOMBRE DE PASSAGERS SUIVANT ASPEM 82 - 83
NOMBRE DE PASSAGERS OBSERVÉS 76

FIGURE 30

CIRCULATION ESTIMEE POUR 82-83 POUR L'O/D YEG-YWG

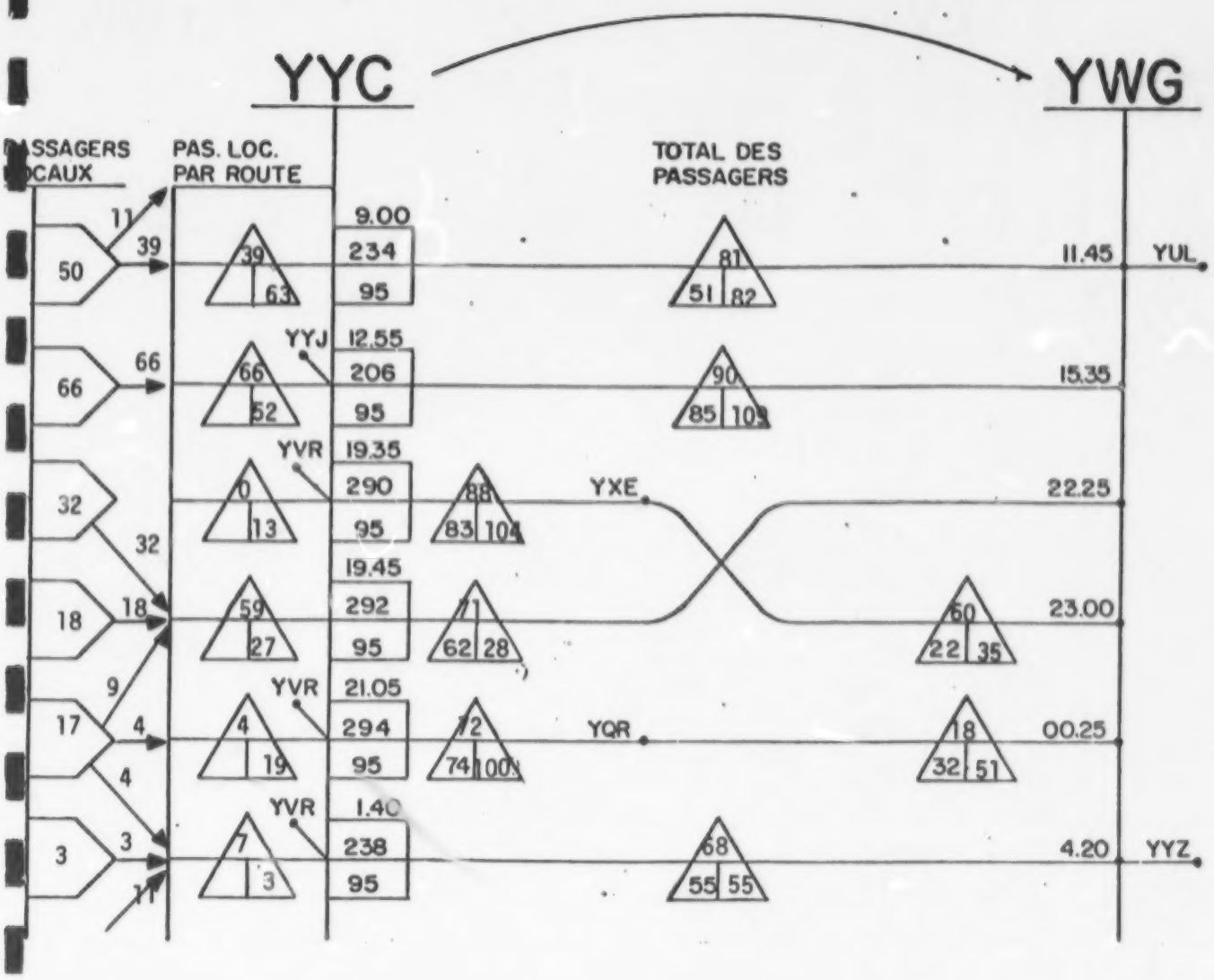
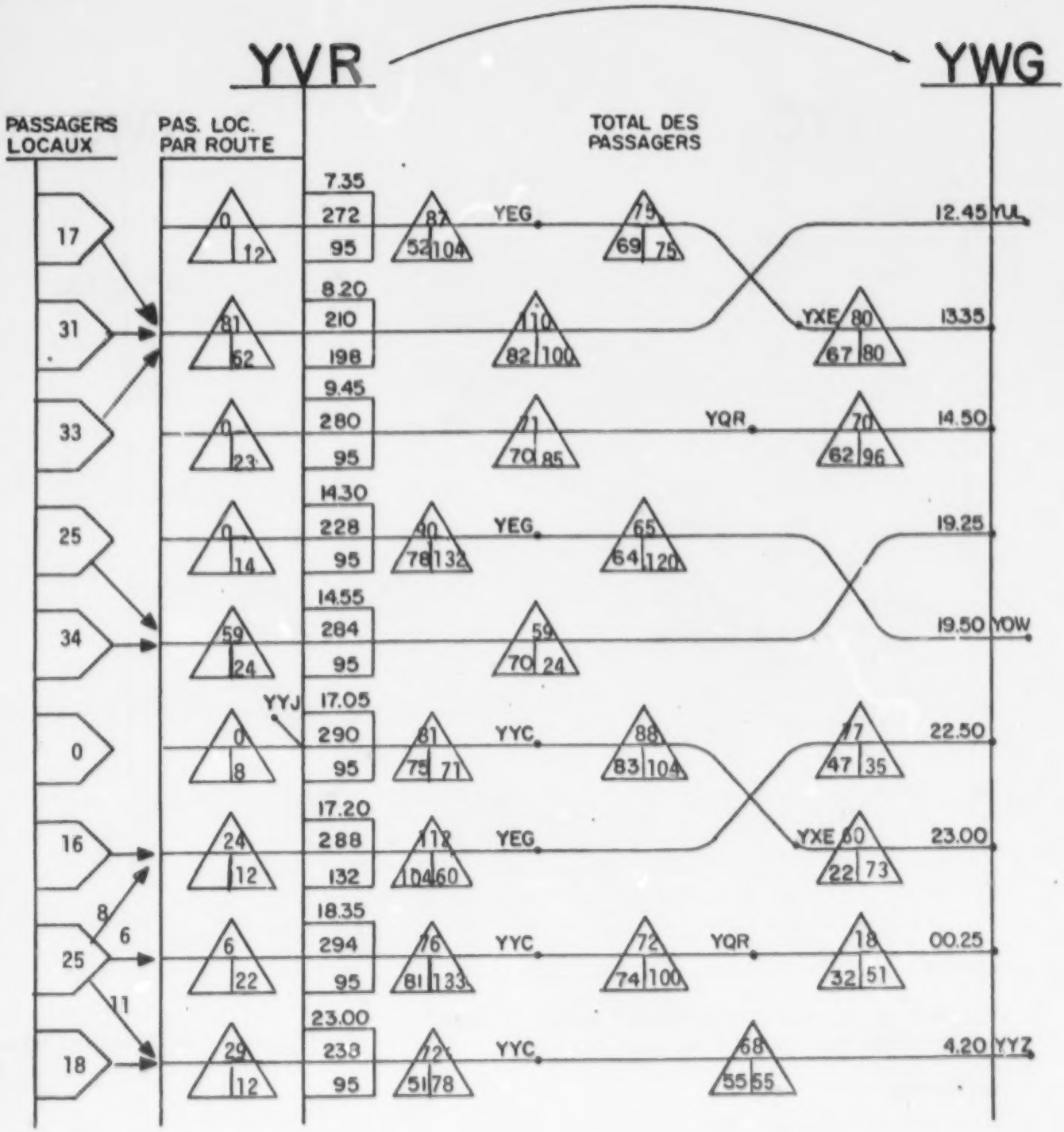



FIGURE 31: CIRCULATION ESTIMEE POUR 82-83 POUR L'O/D YYC-YWG




Coefficient d'utilisation 74%

FIGURE 32: CIRCULATION ESTIMEE POUR 82-83 POUR L'O/D YVR-YWG

CHAPITRE 5

# MODELES DE PLANIFICATION D'UN RESEAU AERIEN

Nous traitons maintenant le problème global en considérant simultanément la circulation des avions et la circulation des passagers. La circulation des avions détermine directement l'horaire, le plan de fréquence et la flotte. Modifier ces variables de décisions correspond à modifier la circulation des avions et inversement. Le problème est donc de modifier la circulation des avions de façon à améliorer la rentabilité des opérations. La circulation des avions détermine les coûts directement; pour connaître les revenus, il faut déterminer la circulation des passagers correspondant à l'horaire associée à cette circulation des avions.

Nous avons proposé dans les chapitres précédents une méthodologie permettant de traiter le problème de circulation des passagers de façon précise et économique. Cette méthodologie va maintenant être utilisée à l'intérieur d'un modèle global avions-passagers pour déterminer l'effet sur les passagers d'une modification de la circulation des avions.

## 5.1 Réseau utilisé par les avions

Avant de discuter le modèle, précisons les notations et le réseau sur lequel se déplacent les avions. Ce réseau est défini à partir de l'ensemble des projets de vols considérés. Ces projets de vols comprennent généralement la plupart des vols existants et d'autres vols que la compagnie est autorisée à offrir et, dont elle veut évaluer le potentiel. L'utilisation de cet ensemble de projets de vols permet de tenir compte des contraintes de nature légale, technique et esthétique limitant la formation des vols.

Le groupe de marketing d'une compagnie devrait pouvoir proposer de tels projets de vols. De plus, des études exploratoires peuvent être réalisées en considérant l'ensemble de vols construit par le système informatique développé par Bourque (76).

Pour chaque type d'avion, un réseau dans l'espace et dans le temps est défini de la façon suivante. Un noeud est défini en chaque ville à chaque heure où un projet de vol débute ou s'y termine. Chaque projet de vol défini un arc en reliant son noeud de départ à son noeud d'arrivée. C'est ce qui distingue le réseau des avions du réseau des passagers où chaque segment de vol entre deux escales forme un arc. Cette différence découle du fait que les passagers peuvent quitter l'avion à chaque escale tandis que le même appareil doit parcourir tout le vol. D'autres arcs sont ajoutés en chaque ville pour relier deux noeuds consécutifs dans le temps. Notons qu'un de ces arcs va du dernier vers le premier noeud de la journée en accord avec le cycle de l'horloge. Ces arcs reliant deux noeuds en une même ville correspondent à des avions attendant au sol. La figure 33 présente un exemple de réseau.

La taille du réseau est réduite par la suite en regroupant les noeuds comme l'indique la figure 33 pour former le réseau agrégé représenté à la figure 34. On groupe des noeuds d'une même ville en les parcourant dans l'ordre chronologique et en débutant un nouveau groupe chaque fois que l'on rencontre un noeud de fin de vol si on a déjà dans le présent groupe un noeud de début de vol. Cette règle d'agrégation assure qu'un vol quittant la ville utilise un avion arrivé précédemment dans cette ville.

Pour apprécier la taille des réseaux à traiter en procédant ainsi, considérons le cas d'Air Canada. Si on ajoute 300 projets de vols aux 325 vols actuels, on obtient 1250 noeuds, donc 1250 arcs de liaisons. Le réseau espace-temps obtenu comprend donc 1250 noeuds et

FIGURE 33: RESEAU ESPACE-TEMPS UTILISE PAR UN TYPE D'AVION

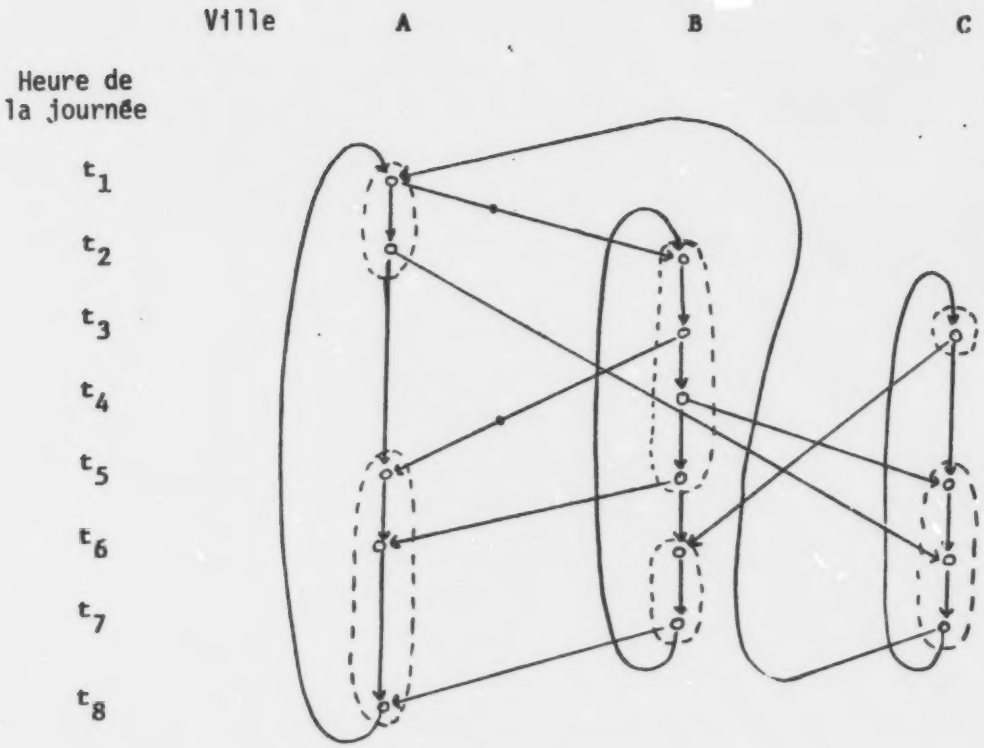
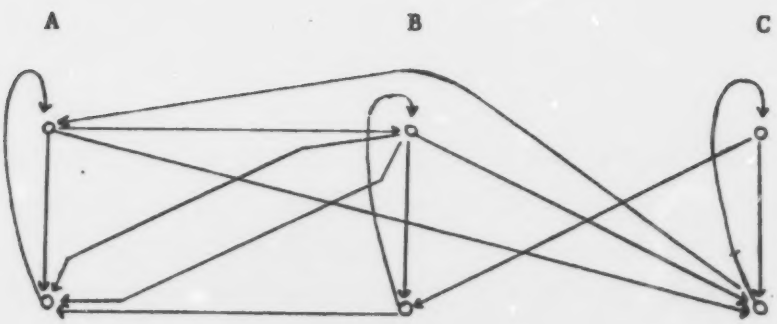



FIGURE 34: RESEAU AGREGE POUR UN TYPE D'AVION



1875 arcs. L'agrégation regroupe les noeuds en groupes moyens de trois, ce qui réduit à 420 noeuds et 1050 arcs. Remarquons de plus, que ce réseau se divise en autant de sous-réseaux qu'il y a de types d'avions. Chez Air Canada, les DC 9 opèrent dans la moitié du réseau et la deuxième moitié se divise entre six autres types d'avions. Même si on ajoutait 5000 projets de vols pour un type d'avions, le réseau réduit serait seulement de l'ordre de 3000 noeuds et 8000 arcs ce qui se traite facilement avec la technologie actuelle.

La capacité des arcs correspondant aux projets de vols est limitée à un. C'est-à-dire qu'au plus un appareil pourra desservir ce vol. Les arcs reliant les différents noeuds d'une même ville n'ont pas de limite de capacité. En effet, plusieurs avions peuvent être en attente en même temps dans la même ville.

La structure de coût est la suivante. Le coût d'opération de chaque projet de vol est associé à l'arc correspondant. Pour intégrer les coûts d'investissements associés à l'achat d'avions, il faut dénombrer les avions utilisés. Ce dénombrement peut être obtenu en comptant les avions sur tous les arcs traversant une heure de référence fixe (minuit par exemple). Ainsi, un coût d'investissement est ajouté à tous les arcs traversant l'heure de référence. Notons que ces coûts peuvent être définis par une compagnie aérienne de façon à intégrer des coûts indirects et d'autres considérations de planification.

## 5.2 Structure des modèles proposés

Avant de discuter le détail de chacun des modèles proposés, voyons leurs caractéristiques communes, et ce qui les distingue des modèles proposés dans la littérature. Cette discussion portera sur la structure de contraintes des présents modèles en comparaison avec

celle des modèles précédents et des conséquences sur la facilité de résolution. Précisons d'abord les notations utilisées.

5.2.1 Notation

Dans la plupart des cas, il s'agit de rappel, ces notations ayant déjà été utilisées précédemment.

Composante passagers

$R$ = l'ensemble des marchés

$d^r$ = la demande pour le marché r

$Q_r$ = l'ensemble des routes desservant le marché r

$h_q$ = le nombre de passagers assignés aur la route q

$h$ = $\{h_q \mid q \in \bigcup_{r \in R} Q_r\}$ une circulation des passagers

$A_q$ = l'attraction de la route q, ($A_q < 0$). $|A_q|$ est la somme pondérée des inconvénients subis en utilisant cette route.

$B^r$ = le coût de refuser un passager du marché r

$L$ = l'ensemble de tous les segments de vols

$\delta_a^q$ = $\begin{cases} 1 \text{ si la route q utilise le segment de vol a} \\ 0 \text{ autrement} \end{cases}$

$R(h, Y)$ = revenu associé à la circulation des passagers

$S(h, Y)$ = mesure de la satisfaction de passagers, exprimée en argent.

Composante avions

$V$ = l'ensemble des projets de vols considérés

$V_k$ = l'ensemble des projets de vols pour les avions du type k

$N$ = l'ensemble des arcs du réseau utilisé par les avions ($V \subset N$)

$y_j$ = le nombre d'avions sur l'arc j. ($_j$ est une variable binaire pour $j \in V$)

$Y$ = $(y_j)$, $j \in N$, la circulation des avions

$\gamma_a^j$ = $\begin{cases} 1 \text{ si le vol j comprend le segment a} \\ 0 \text{ sinon} \end{cases}$

$\Gamma a$ = le nombre maximum de sièges utilisables en moyenne sur le segment de vol a (maximum load factor)

$C_j$ = le coût par avion utilisant l'arc j

$C(Y) = \sum_{J \in N} C_j y_j$ est le coût de la circulation Y

## 5.2.2 Modèles classiques

Voyons maintenant de façon formelle la structure des modèles détaillés discutés à la section 1.2.3 . Les modèles de Agin (72), Chan (74), Richardson (76) et Scarbrough (77) présentent des différences mineures mais ont globalement la même structure. C'est cette structure commune correspondant à l'état de l'art que nous présentons à la

figure 35. Ceci permettra de préciser le type de modifications qu'apportent nos modèles.

Ces modèles tout comme les nôtres considèrent en détail la circulation des passagers et son interaction avec la circulation des avions. L'objectif minimise généralement la différence entre le coût d'opération des avions et le revenu provenant des passagers transportés en considérant une diminution de revenu si la qualité du service n'est pas satisfaisante pour les passagers. Les contraintes (5.1) assurent que la demande de chaque marché est distribuée entre les routes utilisables. Dans un réseau décrit en termes de routes, ces contraintes résument les équations de conservation de flot. Les contraintes (5.2) exigent que les nombres moyens de passagers utilisant chaque avion ne dépassent pas un certain facteur d'utilisation du nombre de sièges. Les équations (5.3) correspondent à la conservation de flot de chaque type d'avions dans le réseau associé. Le bloc (5.4) comprend les contraintes de non-négativité et d'intégrité que doivent satisfaire les variables du problème. Ce problème ne peut être résolu directement pour un grand réseau. Certains auteurs ont essayé de profiter de la structure de bloc des contraintes en utilisant la décomposition de Dantzig-Wolfe. Ils considèrent les contraintes (5.1) et (5.2) comme problèmes maîtres et les blocs de (5.3) comme autant de sous-problèmes. Malheureusement, le problème maître est difficile à résoudre malgré la structure de réseau du problème de circulation des passagers. La présence des contraintes (5.2) en fait un problème de flot à multicommodité avec bornes communes sur les arcs. Ce type de problème est coûteux à résoudre même avec les méthodes les plus avancées (Kennington, 77).

Pour résoudre le problème maître dans un temps raisonnable, on est réduit à utiliser des méthodes plus ou moins heuristiques et les avantages d'un modèle analytique global sont perdus. L'utilisation des contraintes (5.2) pour modéliser l'interaction entre la

FIGURE 35: MODELE CLASSIQUE VS MODELE PROPOSE

MODELE CLASSIQUE

$$\text{Min}\left\{C(Y) - \sum_{r\in R}\sum_{q\in Q_r}(B^r + A_q)\,h_q\right\}$$

(5.1) $\sum_{q\in Q_r} h_q = d^r,\ r \in R$

(5.2) $\sum_{r\in R}\sum_{q\in Q_r} \delta_a^q\, h_q - \sum_k \sum_{j\in V_k} \Gamma_j\, \gamma_a^j\, y_j \le 0,\ a\in L$

(5.3)

Conserv. de flot

Type 1

Type K

(5.4)

| | |
|---|---|
| $h_q \ge 0$ | $y_j \ge 0$, entier, $j\in N$ |
| $q \in Q_r,\ r \in R$ | $y_j = 0$ ou 1, $j \in V$ |

MODELE PROPOSE

$$\text{Min}\left\{C(Y) - R(h) - S(h, Y)\right\}$$

$\sum_{q\in Q_r} h_q = d^r,\ r \in R$ — Circulation des passagers

Limite de capacité des avions

Conserv. de flot

Type 1

Type K

Circulation des avions

| | |
|---|---|
| $h_q \ge 0$ | $y_j \ge 0$, entier, $j\in N$ |
| $q \in Q_r,\ r \in R$ | $y_j = 0$ ou 1, $j \in V$ |

circulation des passagers et celle des avions est à la source des difficultés de résolution. De plus, ce type de contraintes représente assez mal la limite de capacité des avions. Ces contraintes sont le point faible de ces modèles et doivent être remplacées.

### 5.2.3 Modèles proposés

La conclusion que dégage clairement la section précédente, c'est qu'il faut remplacer les contraintes (5.2). Plutôt que de les remplacer, nous avons choisi de les faire disparaître, en vertu des considérations suivantes. Même si ce n'est pas souhaitable, il n'est pas illégal ou impossible d'offrir moins de sièges qu'il y a de passagers à transporter. Il n'y a donc pas lieu de mettre une contrainte stricte entre le nombre de passagers et la capacité de l'avion. Par contre, il faut tenir compte des conséquences économiques d'une telle situation et réajuster les estimations de revenus en tenant compte des passagers qui ne sont pas transportés, faute de place. C'est donc au niveau de l'objectif que doit intervenir l'interaction de la circulation des passagers avec la circulation des avions.

Ainsi, comme l'indique la figure 35, nous considérons comme réalisable tout couple (h, Y) satisfaisant (5.1), (5.3) et (5.4). Il suffit qu'un tel couple soit composé d'une circulation des passagers et d'une circulation des avions, la compatibilité entre ces deux circulations est mesurée par l'objectif. Ceci donne beaucoup de souplesse pour la modification de la circulation des avions; on peut retirer temporairement les avions desservant une paire de villes s'il y a une façon plus rentable de les utiliser ailleurs et laisser ces villes sans service en attendant d'y déplacer des avions mieux adaptés.

Du point de vue de la difficulté de résolution, la disposition des contraintes (5.2) améliore grandement la structure du problème.

Les contraintes se résument maintenant à des blocs indépendants de contraintes de conservation de flot; un bloc par marché pour les variables de passagers, et un bloc par type d'avions. L'interaction entre ces blocs intervient seulement au niveau de la fonction économique.

## 5.3 Formulation des modèles proposés

Au chapitre 2, nous avons développé quatre méthodes pour déterminer la circulation des passagers sans utiliser de contraintes explicites de capacité des avions. Ces méthodes utilisent la notion de passagers excédentaires pour déterminer la perte de revenus dans l'objectif ou pour calculer les réallocations de passagers dans les conditions d'équilibre.

Chacune de ces méthodes permet de définir un modèle de planification de flotte du type proposé. Il suffit de lui adjoindre les variables Y, les contraintes de circulation d'avions et de définir l'objectif global incluant le coût d'utilisation des avions. Ces quatre modèles sont appelés G1, G2, G3 et G4 en correspondance avec P1, P2, P3 et P4, les problèmes de circulation des passagers sous-jacents.

### 5.3.1 Modèles G1 et G2: planification d'un réseau aérien en traitant la circulation des passagers suivant l'optimisation coopérative

Les contraintes ayant été définies, précisons maintenant l'objectif de chacun de ces modèles. Les objectifs sont de la forme

$$\text{Min } \{C(Y) - R(h, Y) - S(h, Y)\} \qquad (5.5)$$

Le coût de la circulation des avions est donné par $C(Y) = \sum_{j \in N} C_j$ .

Le revenu s'obtient en prenant la différence entre le coût total des billets de clients voulant voyager et le revenu perdu en refusant des passagers, faute de place. Le revenu varie comme $- Z_1$, le revenu perdu.

Les deux modèles se distinguent par la façon de définir $S(h, Y)$, la mesure de la satisfaction. Pour le modèle G1, $S_1(h, Y)$ est défini par l'équation (2.14)

$$S_1(h_1, Y) = -KZ_2 = K \sum_r C^r \sum_{q \in Q_r} h_q A_q$$

$A_q$ est le négatif de la somme pondérée des inconvénients de la route q. Ainsi, $S_1(h, Y)$ est le négatif de la somme pondérée des inconvénients subis par tous les passagers. L'objectif du premier modèle est à une constante près

$$\text{Min } \{C(Y) + \qquad (5.6)$$

C'est précisément l'objectif du modèle P1 auquel on ajoute le coût de la circulation des avions.

Pour le modèle G2, l'objectif est encore de la forme (5.6) mais, dans ce cas, $Z_2$ est défini par la méthode développée à la section (2.5.3).

$$Z_2(h) = \sum_r \sum_{q \in Q_r} \int_0^{h_q} g\left(\frac{x}{A_q}\right) d\,x$$

Dans ce cas, $Z_2$ mesure le coût de s'écarter de la répartition en proportion des $A_q$ .

### 5.3.2 Modèle G3: planification d'un réseau aérien en traitant la circulation des passagers suivant l'optimisation égoïste

Le modèle P3 ne détermine pas la circulation des passagers en fonction d'un objectif mais plutôt en utilisant des conditions d'équilibre. L'attraction des routes et le calcul des passagers excédentaires interviennent pour déterminer les réallocations nécessaires et la circulation des passagers. Ce modèle considère que tous les passagers voyagent; il n'y a donc pas de variations de revenus. Ainsi, il n'a pas de termes mesurant le revenu et la satisfaction pour faire contrepoids au coût de circulation des avions.

Un objectif peut être formé en introduisant $S_3(h, Y)$ une mesure de satisfaction de l'horaire. L'objectif devient

$$\text{Min } \{C(Y) - S_3(h, Y)\} \tag{5.7}$$

Remarquons toutefois que h n'est pas une variable indépendante puisque Y détermine h via les conditions d'équilibre. L'objectif doit être traité comme une fonction de Y seulement.

### 5.3.3 Modèle G4: planification d'un réseau aérien en considérant une demande variable et la circulation des passagers suivant l'optimisation égoïste

Dans ce cas, la demande et le revenu varient avec le niveau de service. L'objectif suivant

$$\text{Min } \{C(Y) - R(h, Y)\} \tag{5.8}$$

est une bonne mesure de l'intérêt de la compagnie de modifier la circulation des avions. Il n'est pas nécessaire d'introduire de termes mesurant la satisfaction des passagers. L'insatisfaction se reflète sur le revenu.

Précisons maintenant le calcul du revenu. Comme nous l'avons discuté à la section (2.7), le nombre de passagers transportés est fonction de l'attraction et de la disponibilité des routes offertes. Le nombre de clients $d^r$ attirés par le marché r dépend de l'attraction des routes offertes. Donc, il doit s'écrire $d^r(Y)$. La partie de ces clients $v^r$ qui voyagent effectivement est fonction de la disponibilité des routes qui elle dépend de la circulation des passagers. $v^r$ est déterminé en fonction de $d^r(Y)$ par le modèle P4 en même temps que la circulation des passagers.

Pour déterminer le revenu associé à Y, il reste à préciser la fonction $d^r(Y)$. Une telle fonction a été proposée par Eriksen and Teneja (75) qui expriment la demande en fonction d'un indice de qualité de service, des tarifs et du pouvoir d'achat dans ce marché à l'aide d'un modèle multiplicatif. Un autre travail réalisé par Cambridge Systematics (75) propose d'utiliser le modèle LOGIT pour déterminer le nombre total des passagers attirés par chaque vol.

## 5.4 Méthodes de résolutions

Considérons d'abord les modèles G1 et G2 où h et y sont déterminés en fonction d'un objectif global. Ces problèmes sont de la forme

$$\begin{aligned} &\text{Min } \{C(Y) + Z(h, Y)\} \\ &\text{s.a. } h \in H(Y) \\ &\quad Y \in \mathcal{Y} \end{aligned} \tag{5.9}$$

où Z(h, Y) correspond à l'objectif des modèles P1 ou P2 suivant le cas. H(Y) est l'ensemble des circulations des passagers dans le

réseau défini par l'horaire associé à Y. $y$ est l'ensemble des circulations des avions.

Pour résoudre ce problème, nous le considérons plutôt sous la forme suivante

$$\text{Min } \{C(Y) + \min_{h \in H(Y)} Z(h, Y)\}$$

$$\text{s.a.} \quad Y \in y \tag{5.10}$$

Seul Y demeure comme variable de décision, le calcul de h n'est qu'une étape intermédiaire pour faire évaluer l'objectif. Ceci correspond exactement au rôle joué par ces variables dans le processus de planification.

Ce problème est résolu en utilisant une adaptation de la méthode de Frank-Wolfe. Revoyons maintenant cette méthode avant d'expliquer comment nous l'avons modifiée pour résoudre ce type de problème.

## 5.4.1 Méthode de Frank-Wolfe

La méthode de Frank-Wolfe résout le problème

$$\text{Min } W(Y)$$

$$\text{s.a. } Y \in y \tag{5.11}$$

de façon itérative en se déplacant d'une solution réalisable $Y^k$ à une autre solution réalisable $Y^{k+1}$ tel que $W(Y^{k+1}) < W(Y^k)$. De plus, $Y^{k+1}$ est déterminé à partir de $Y^k$ par la relation

$$Y^{k+1} = Y^k + \sigma_k \, d^k$$

où $d^k$ est une direction de descente et $\sigma_k$ le déplacement optimal dans

cette direction. La direction $d^k$ est celle qui permet la plus grande descente de l'approximation linéaire de la fonction objective. Elle est solution du problème suivant

$$\begin{aligned} &\text{Min} \quad \nabla W(Y^k) \cdot (Y^k + d^k) \\ &\text{s.a.} \quad Y^k + d^k \in \mathcal{Y} \end{aligned} \tag{5.12}$$

Le déplacement $\sigma_k$ est choisi de façon à minimiser W en se déplaçant dans la direction $d^k$

$$W(Y^k + \sigma_k d^k) = \min_{0 \le \sigma \le 1} W(Y^k + \sigma d^k) \tag{5.13}$$

### 5.4.2 Modifications apportées

Deux modifications de la méthode de Frank-Wolfe sont nécessaires pour tenir compte de la structure du problème. Premièrement, les variables $y_j$ doivent prendre seulement des valeurs entières car elles représentent des nombres d'avions. Deuxièmement, la fonction objective de (5.10) n'est pas définie explicitement; elle contient la résolution du problème de circulation des passagers.

Rappelons que l'ensemble $\mathcal{Y}$ est défini par des contraintes de conservation de flot. Il est bien connu qu'un problème ayant un objectif linéaire et de telles contraintes a une solution optimale dont les variables sont entières. Comme la méthode de Frank-Wolfe résout des problèmes du type (5.12) ayant un objectif linéaire, la solution $Y^{k+1}$ reste entière si $\sigma_k$ est égal à 0 ou 1. Nous utiliserons donc seulement les déplacements entiers pour conserver Y entier

Le déplacement est calculé ainsi:

Si $W(Y^k + d^k) < W(Y^k)$ alors $\sigma_k = 1$, sinon $\sigma_k = 0$ .

Ceci correspond à déterminer $Y^k + d^k$ une circulation alternative des avions en résolvant une linéarisation du problème. Ensuite, on évalue l'objectif pour cette circulation alternative en résolvant le problème des passagers pour l'horaire associé. Si l'objectif est amélioré, on conserve cette solution alternative, sinon, on conserve la même solution.

Cette approche présente des similarités avec la "reflection programming" proposée par Agin, (74). Remarquons que cette méthode est sous-optimale et peut terminer avec un optimum local, mais elle donne des solutions entières. Par contre, il ne semble pas possible de résoudre de façon optimale en nombres entiers ce problème de plusieurs milliers de variables. Minoux, (75) qui a utilisé une méthode du même type pour traiter un réseau de communication rapporte des résultats satisfaisants du point de vue de l'optimalité.

La méthode de linéarisation de l'objectif doit être adaptée pour traiter cet objectif défini implicitement. Rappelons que le gradient est utilisé comme linéarisation pour la méthode standard. A l'itération k, il faut calculer

$$\frac{\partial}{\partial y_j}\Big(c(Y) + \underset{h\in H(Y)}{\text{Min}} \; Z(h, Y)\Big)\Big|_{Y=Y^k} = C_j + \frac{\partial}{\partial y_j}\Big(\underset{h\in H(Y)}{\text{Min}} \; Z(h, Y)\Big)\Big|_{Y= Y^k}$$

Posons $\bar{Z}_j^k = \frac{\partial}{\partial y_j}\Big(\underset{h\in H(Y)}{\text{Min}} \; Z(h, Y)\Big)\Big|_{Y = Y^k}$

Pour les arcs $j \in N - V$ (arcs d'attente)

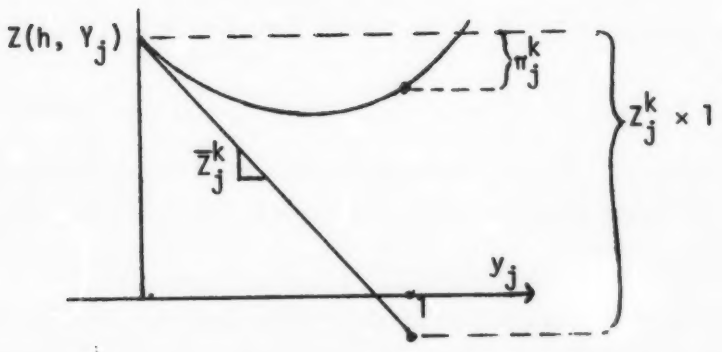
$\bar{Z}_j^k = 0$ car, aucun passager n'utilise ces arcs. Pour un arc $j \in V$ (vols) comme la modification correspondant à ajouter ou enlever le vol, il est préférable d'utiliser la différence finie plutôt que la dérivée pour évaluer la modification de revenu associé. En prenant $\Delta_j$ égal au

au vecteur unité $e_j(-e_j)$ si le vol inactif (actif) j est ajouté (discontinué), la différence finie $\pi_j^k$ s'évalue ainsi

$$\pi_j^k = \underset{h\epsilon H(Y^k+\Delta_j)}{\text{Min}} Z(h, Y^k+\Delta_j) - \underset{h\epsilon H(Y)}{\text{Min}} Z(h,Y) \qquad (5.14)$$

La figure 36 illustre l'avantage de la différence finie par rapport à la dérivée pour évaluer la modification de revenus.

FIGURE 36: Linéarisation de la fonction objective avec les différences finies




La direction de descente de la méthode de Frank-Wolfe modifiée est donc la solution optimale du problème suivant

$$\begin{aligned} &\text{Min} \quad (C + \pi^k)\,(Y^k + d^k) \\ &\text{s.a.} \quad Y^k + d^k \in Y \end{aligned} \qquad (5.15)$$

(en posant $\pi_j^k = 0$ pour $j \in N - V$)

Pour évaluer $\pi_j^k$, il faut résoudre le problème de circulation des passagers pour l'horaire associé à $Y^k + \Delta_j$ . Pour réduire le temps de calcul, les valeurs des $\pi_j^k$ sont estimées de la façon suivante.

S'il s'agit d'enlever un vol présent, certaines routes de certains marchés ne sont plus disponibles. En traitant un à un les marchés touchés, on réassigne les passagers de ce vol sur des routes alternatives et $\pi_j^k$ est défini comme le total des augmentations de coût résultantes. Cette méthode surestime légèrement les coûts car, elle ne permet pas le réarrangement des passagers des vols non touchés.

S'il s'agit d'ajouter un nouveau vol, de nouvelles routes sont définies pour certains marchés. En traitant ensemble ces marchés touchés, on détermine la circulation des passagers tout en maintenant fixés les choix de routes des passagers des marchés non touchés. $\pi_j^k$ est défini comme le négatif de la diminution de coûts résultante.

Ces valeurs de $\pi_j^k$ négligent les interactions entre les modifications de vols. Ainsi, lors du calcul de $d^k$ en résolvant (5.15), les $C_j + \pi_j^k$ évaluent bien l'impact financier des premières modifications mais ne tiennent pas compte de l'effet de ces premières modifications sur les bénifices attendus des modifications subséquentes. Les solutions alternatives comportant beaucoup de modifications risquent donc de ne pas améliorer l'objectif car, elles reposent sur des estimations de bénifice qui ne sont pas valables. Il est donc préférable de ne pas résoudre (5.15) jusqu'à l'optimalité mais, de déterminer une solution alternative ne comportant que quelques modifications. On a ainsi de meilleures chances que cette solution alternative augmente le bénifice.

5.4.3 Algorithme

Résumons maintenant les étapes de cette méthode.

Etape 0

Trouver $Y^0$ une première circulation des avions (la circulation actuelle peut entre autres être utilisée). Déterminer, en utilisant l'un des modèles proposés, la circulation des passagers correspondant à l'horaire associé.

Etape 1

Evaluer les différences finies $\pi_j^k$ pour chaque vol proposé.

Etape 2

Résoudre (pas nécessairement jusqu'à l'optimalité) le problème (5.15) pour déterminer $Y^k + d^k$ une circulation alternative des avions.

Etape 3

Evaluer la valeur de l'objectif à $Y^k + d^k$ . Ceci demande de déterminer la circulation des passagers correspondant à l'horaire associé à $Y^k + d^k$ .

Etape 4

Si $Y^k + d^k$ diminue la valeur de l'objectif i.e.

$$C(Y^k + d^k) + \min_{h \in H(Y^k + d^k)} Z(h, Y^k + d^k) \leq C(Y^k) + \min_{h\ H(Y)} Z(h, Y^k)$$

on prend $Y^{k+1} = Y^k + d^k$ et retourne à l'étape 1.

Etape 5

Si $Y^k + d^k$ ne diminue pas la valeur de l'objectif, on retourne à l'étape 2 pour déterminer une autre circulation alternative.

On continue cette procédure tant qu'il est possible de trouver des circulations alternatives susceptibles de diminuer la valeur de l'objectif.

Plusieurs stratégies de solution peuvent être utilisées. Par exemple, on pourrait traiter les types d'avions un à la fois et ainsi, à l'étape 1, évaluer seulement les $\pi_j^k$ des vols utilisant ce type d'avions et à l'étape 2, déterminer $d^k$ en résolvant seulement le sous-problème de circulation de ce type d'avions. De même à l'étape 2, on peut déterminer des solutions alternatives s'écartant plus ou moins de la solution courante. Des essais numériques seront nécessaires pour choisir les meilleures stratégies.

Les techniques de solutions utilisées à chacune des étapes sont très efficaces. Le tableau 8 donne un aperçu du temps par itération.

TABLEAU 8: Temps de résolution du problème global

Initiation

| | |
|---|---|
| - Construction du réseau | 30(*) |
| - Etape 0: Déterminer la circulation des passagers | 20 |
| - Temps d'initialisation | 50 |

(*) Temps en secondes de CPU d'un CDC 173.

Itérations suivantes

| | |
|---|---|
| - Etape 1: Calcul des $\pi_j^k$ | 5 |
| - Etape 2: Déterminer une circulation alternative des avions | 3 |
| - Mise à jour du réseau des passagers | 10 |
| - Etape 3: Réoptimisation de la circulation des passagers | 8 |
| - Temps par itération | 26 |

L'étape 0 et l'étape 3 sont réalisées avec MAPUM qui est très rapide, comme nous l'avons vu au chapitre 4. Le temps de l'étape 3 est moindre que les 20 secondes de l'étape 0 car on débute avec une très bonne solution et il suffit de réoptimiser. L'étape 1 utilise la même technologie pour résoudre les mini problèmes lors du calcul des $\pi_j^k$. L'étape 2 réalise quelques modifications de flot dans un réseau de taille moyenne. En utilisant l'algorithme OUT-OF-KILTER, on a pu réaliser 100 modifications de flot en moins de trois secondes dans un réseau de 300 noeuds et 1000 arcs.

Avec de tels temps, on peut réaliser 20 itérations en dix minutes dans un réseau de 300 vols. Ceci devrait permettre de modifier un horaire de façon appréciable dans un temps raisonnable.

### 5.4.4 Adaptation de la méthode de résolution pour les modèles G3 et G4

Ces modèles présentent les mêmes contraintes que les précédents. Leurs objectifs sont encore composés de deux termes, C(Y) le coût des avions et un deuxième terme dépendant de la circulation des passagers.

La différence, par rapport aux modèles précédents, est que ce deuxième terme n'est plus l'objectif du problème de circulation des passagers. De plus, les variables h ne concourent pas à minimiser l'objectif global car, pour un Y donné, elles sont complètement déterminées par des conditions d'équilibre.

Pour ces deux modèles, les problèmes à traiter sont du type suivant

$$\text{Min } \{ C(Y) + T(h, Y) \}$$
$$Y \in \mathcal{Y} \qquad (5.16)$$

où h est obtenue en fonction de Y par la résolution du problème de circulation des passagers et T(h, Y) représente S(h, Y) ou R(h, Y) suivant qu'il s'agit du modèle G3 ou G4.

Le problème (5.16) est résolu tout comme précédemment par la méthode de Frank-Wolfe. La linéarisation de l'objectif pour déterminer la direction d'amélioration utilise encore les différences finies. Dans ce cas, il faut calculer la variation de T(h, Y) pour une variation $\Delta_j$ de l'horaire. $\Delta_j$ égale le vecteur unité $e_j$ ($-e_j$) si le vol inactif (actif) j est ajouté (discontinué)

$$\pi_j^k = T(h^1, Y^k + \Delta_j) - T(h^2, Y^k) \qquad (5.17)$$

où $h^1$ et $h^2$ correspondent respectivement à la circulation des passagers pour l'horaire associé à $Y^k + \Delta_j$ et à $Y^k$.

Ces différences finies sont estimées tout comme précédemment en résolvant de petits problèmes d'allocation des passagers. L'algorithme de résolution est analogue au cas précédent et devrait permettre d'améliorer la circulation des avions dans un grand réseau et ce dans un temps de calcul très raisonnable.

## CHAPITRE 6

# CONCLUSION

Ce travail représente un pas important pour la planification de flotte et l'organisation d'horaires en tenant compte des interactions complexes entre la circulation des passagers et celle des avions. La notion de "passagers excédentaires" permet un traitement plus réaliste de cette interaction. De plus, elle conduit à des structures de modèles se traitant avec des méthodes de décomposition et des techniques de programmation mathématique très efficaces. Ceci permet de traiter de grands réseaux avec un haut niveau de détail dans des temps de calcul très courts.

Une grande partie de ce travail porte sur la circulation des passagers, un point faible des modèles précédents. Quatre modèles sont proposés. Ils introduisent plusieurs nouveaux concepts permettant de représenter plus fidèlement la réalité.

D'abord, un modèle simpliste dont la seule qualité est d'utiliser la notion de "passagers excédentaires" pour modéliser la limite de capacité des avions. La précision des résultats obtenus malgré le traitement simpliste des autres facteurs, montre l'importance de cette notion.

Le deuxième modèle introduit en plus, le comportement pluraliste des passagers. Plutôt que de tous tenter d'utiliser la route la plus attrayante, les passagers tentent de se reporter entre les routes en proportion de l'attraction.

Le troisième modèle reproduit en plus, le comportement égoïste des passagers (user optimisation). Chaque passager choisit son

itinéraire en fonction de sa seule satisfaction sans se préoccuper de celle des autres ni de l'utilisation rationnelle du réseau.

Le dernier modèle intègre en plus, l'équilibre offre/demande. La demande varie avec le niveau de service, qui est fonction de la disponibilité des routes qui dépend de la circulation des passagers.

Le premier de ces modèles a été expérimenté dans un grand réseau pour une saison passée et, comparé avec la circulation observée. La moyenne des valeurs absolues des erreurs d'estimation par arc, est de 11.3 passagers. Ceci est considéré comme une bonne précision. La version du modèle ASPEM utilisé par Air Canada a produit une erreur deux fois plus grande dans les mêmes conditions.

Deux sources d'erreurs expliquent les différences, les nombres de passagers estimés par le modèle et les nombres de passagers observés. Une première source est l'imprécision des données d'entrée. Par exemple,les distributions d'heures de départ désirées étaient particulièrement imprécises. Le tableau 7 illustre des écarts de 20% et même 30% entre les nombres de passagers observés par demi-journée et les nombres calculés suivant ces courbes. De même, la méthode d'estimation des écarts-types de la demande par arc pourrait être améliorée. En remplaçant les écarts-types estimés par ceux observés, l'erreur moyenne peut être réduite de 11.3 à 9.1 . Une seconde source d'erreur provient du fait que ce premier modèle néglige certains aspects de la réalité. En introduisant dans le modèle le traitement des particularités tel l'attraction des gros avions, la priorité d'utilisation du service Rapidair, on pourrait corriger les plus grosses erreurs. De plus, les résultats devraient être améliorés en utilisant les derniers modèles qui considèrent un comportement plus réaliste des passagers face à l'attraction et à la disponibilité des routes.

Les méthodes de résolution ont reçu une attention toute particulière. Des techniques ont été développées pour profiter au maximum des structures particulières des problèmes. Les développements théoriques nécessaires pour montrer l'existence de solutions et la convergence des méthodes proposées ont été réalisés. Ceci permet d'obtenir des solutions optimales à coût minime. Rappelons que le modèle P1 a déterminé la circulation des passagers dans un réseau de 76 villes et 325 vols en 20 secondes CPU (CDC 173).

Avec un temps de résolution aussi court, un tel modèle de répartition des passagers peut être utilisé de façon interactive par le planificateur pour évaluer et modifier un horaire. Il suffirait d'ajouter des modules d'entrée/sortie pour compléter un tel système. L'implantation de ce système interactif devrait être facilitée par le fait qu'il laisse l'initiative aux planificateurs; il ne fait que leur donner un outil pour quantifier l'impact de leurs projets de modification d'horaire. Ce court temps de résolution est aussi une qualité essentielle pour l'intégration d'un modèle de répartition de passagers, dans un modèle de planification de flotte. Il permet de réaliser plusieurs modifications de la circulation des avions dans un temps raisonnable.

Finalement, soulignons une innovation méthodologique qui devrait avoir des retombées sur la modélisation de la circulation automobiles et du transport en commun. Les modèles d'équilibre reproduisant l'optimation égoïste ne se ramènent pas, en général, à des problèmes de programmation mathématique. Ils donnent lieu à des problèmes de complémentarité non-linéaire très coûteux à résoudre quand on utilise des hypothèses de comportement extrémistes tel que le choix de la route la plus attrayante, la plus courte ou la moins coûteuse. En utilisant des hypothèses de comportements pluralistes impliquant une dispersion des passagers entre toutes les routes utilisables comme nous l'avons fait, on obtient des systèmes d'équations non-linéaires qui peuvent être résolues efficacement par la méthode de Gauss-Siedel.

L'utilisation de la notion de "passagers excédentaires" pour traiter la limite de capacité des avions permet de formuler des modèles de planification de flotte et d'horaire où l'interaction entre la circulation des passagers et celle des avions intervient seulement au niveau de l'objectif. Cette structure permet la résolution par décomposition en traitant alternativement la composante passagers et la composante avions.

L'adaptation de la méthode de Frank-Wolfe utilisée pour résoudre ce problème permet d'obtenir des solutions entières pour la circulation des avions mais n'est pas nécessairement optimale. Elle permet toutefois d'améliorer une solution réalisable en considérant l'impact sur les coûts et les revenus dans l'ensemble du réseau. C'est probablement le mieux que l'on peut espérer pour l'instant car il n'est pas possible de résoudre de façon optimale en nombres entiers ce problème non-linéaire de plusieurs milliers de variables.

Beaucoup de travail reste à faire avant d'en arriver à un "logiciel" général de planification d'un réseau aérien. Par contre, la technologie introduite permet des gains importants du point de vue des propriétés théoriques, de la validité des résultats et du temps de calcul. Il devient possible de formuler un modèle global traitant l'essentiel d'un grand problème dans un temps raisonnable. Il reste à développer les méthodes d'estimation des paramètres d'entrée, à calibrer ces modèles et à faire la démonstration de leur utilité dans un éventail de situations pratiques.

## ANNEXE 1

### Validation des hypothèses utilisées

Nous discutons ici la valeur des deux hypothèses utilisées lors du calcul du nombre de passagers excédentaires. La validation est faite avec les données de l'hiver 1976 pour l'ensemble du réseau d'Air Canada. Dans ce cas, les résultats sont très satisfaisants mais il faudra étudier d'autres périodes de l'année et d'autres réseaux avant d'accepter ces hypothèses comme résultat général et définitif.

Si ces hypothèses devaient être modifiées pour représenter d'autres situations, le modèle pourrait être adapté facilement. Ainsi, la présente discussion se limite à montrer que les hypothèses permettent d'obtenir des résultats approchant la réalité et n'a pas la prétention de montrer que c'est la seule bonne façon de traiter le problème.

Rappelons les deux hypothèses utilisées:

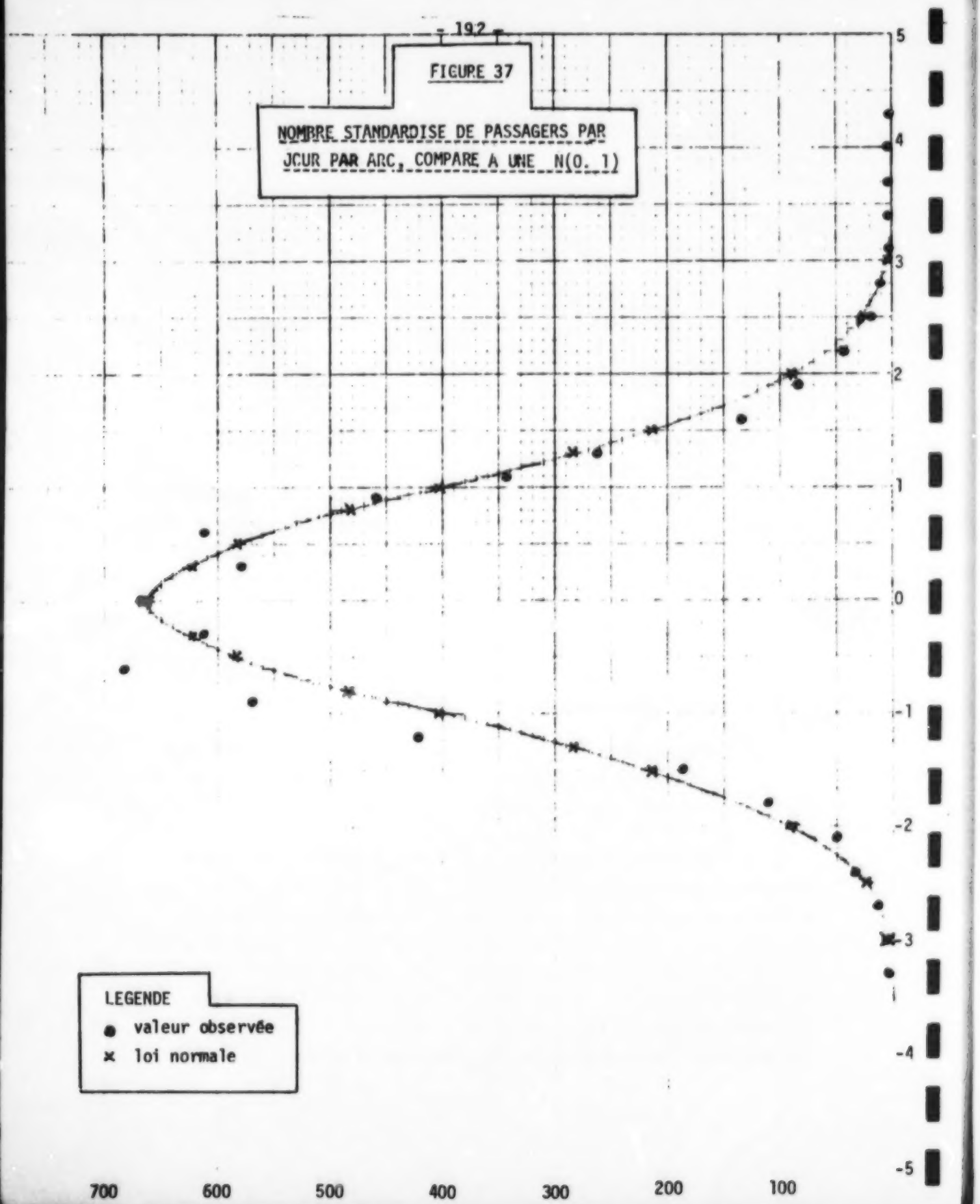
A) Le nombre journalier de passagers suit approximativement une distribution normale.

B) Les nombres journaliers de passagers de deux points O-D sont si fortement corrélés que l'erreur n'est pas très grande si on utilise 1 comme coefficient de corrélation.

L'hypothèse de normalité a d'abord été étudiée pour chaque arc. Des tests de Kolmagarov-Smirnov à 90% ont été faits sur la normalité de la distribution des 89 nombres journaliers de passagers (3 mois) sur chaque arc. L'hypothèse de normalité a été acceptée pour 321 sur 536 arcs. La présence de nombres de passagers discordants (outliers), autant très grands que très petits, est une des causes fréquentes de rejet. Une autre cause de rejet est la présence de créneaux trop prononcés dans les distributions observées. Par contre,

FIGURE 37

NOMBRE STANDARDISE DE PASSAGERS PAR JOUR PAR ARC, COMPARE A UNE N(0. 1)



les distributions non acceptées comme normales ne se ressemblent pas entre elles. Elles ne présentent ni étirement systématique à gauche ou à droite, ni aplatissement.

La deuxième étude regroupe les données de tous les arcs, ce qui permet de neutraliser les perturbations dues aux petits échantillons et donne une mesure globale de normalité. Les données une fois standardisées avec la moyenne et l'écart-type de chaque arc peuvent être regroupées, car elles deviennent toutes N(0,1) si l'hypothèse de normalité est vraie. Nous obtenons ainsi un échantillon de journées-arcs. La figure 37 représente cet échantillon composé avec une normale (0,1). Du point de vue des tests statistiques de normalité, cette distribution est symétrique avec un moment d'ordre trois de 0.007. L'aplatissement est de 2.91, ce qui est acceptable. Par contre, elle n'est pas acceptée comme normale par un test de Kalmogorov-Smirnov à 90%. Ses défauts entre -1 et 0 sont trop importants. Or, comme seulement la partie droite de la courbe est utilisée par le modèle et que cette partie s'approche davantage d'une normale, nous avons fait pour la partie droite de la courbe un test de Kalmogorov-Smirnov qui accepte la normalité à 90%. Par contre, un test $\chi^2$ à 90% rejette la normalité.

Comme nous le constatons, les tests statistiques ne sont pas unanimes sur la normalité; les résultats varient suivant le type de défauts auxquels les tests sont sensibles. Pour étudier la normalité de façon pertinente pour l'application, nous avons évalué la différence entre l'espérance du nombre de passagers excédentaires calculé en utilisant la loi normale et en utilisant la distribution empirique. Le tableau 2 présente ces résultats. La partie gauche du tableau présente la comparaison dans le cas de distributions standardisées. La partie droite présente un exemple type:

$$C = 100 \text{ places} \quad , \quad \sigma_a = \lambda \, . \, f_a = .4 \, f_a \quad .(*)$$

---

(*) Dans le réseau domestique américain la valeur moyenne du coefficient de variation est $\lambda = .38$ .

TABLEAU 8

DIFFERENCE ENTRE LES NOMBRES DE PASSAGERS EXCEDENTAIRES CALCULES EN UTILISANT LA LOI NORMALE ET EN UTILISANT LA DISTRIBUTION EMPIRIQUE

| EN PASSAGERS STANDARDISES | | | EXEMPLE: C = 100. λ = 4 | | |
|---|---|---|---|---|---|
| (1) p | (2) E(p,0,1) | (3) différence | (4) $f_a$ | (5) passagers excédentaires | (6) différence |
| -.25 | .534 | .000 | 111 | 23.7 | 0 |
| -.15 | .478 | +.011 | 106 | 20.0 | +.5 |
| -.05 | .399 | +.027 | 102 | 16.0 | +1.0 |
| +.05 | .374 | -.025 | 98 | 15.0 | -1.0 |
| +.15 | .328 | +.031 | 94 | 12.4 | +1.1 |
| +.25 | .287 | +.028 | 91 | 10.4 | +1.0 |
| +.35 | .248 | +.027 | 88 | 8.7 | +.7 |
| +.45 | .214 | +.006 | 85 | 7.2 | +.2 |
| +.55 | .183 | -.005 | 82 | 6.0 | -.2 |
| +.65 | .155 | -.022 | 79 | 4.9 | -.7 |
| +.75 | .131 | -.023 | 77 | 4.0 | -.7 |
| +.85 | .110 | -.029 | 75 | 3.3 | -.8 |
| +.95 | .092 | -.023 | 72 | 2.6 | -.7 |
| +1.15 | .062 | -.011 | 70 | 1.7 | -.3 |
| +1.35 | .041 | -.001 | 68 | 1.1 | -.03 |
| +1.55 | .026 | .003 | 67 | .6 | +.07 |
| +1.75 | .016 | .013 | 65 | .4 | +.3 |
| +1.95 | .010 | .009 | 63 | .2 | +.2 |

LEGENDE:

(1) Places disponibles standardisées $p = (c - f_a)/\sigma_a$
(2) Passagers refusés en mesure standardisée avec la loi normale
(3) E(p, 0, 1) mais l'estimation utilisant la distribution empirique
(4) Demande moyenne correspondant au p de cette ligne
(5) Passager refusé estime avec la loi normale
(6) Valeur (5) moins l'estimation utilisant la distribution empirique

Dans ce cas, la différence est en nombre de passagers. Ce tableau justifie l'utilisation de la loi normale. Les différences sont au plus de l'ordre de un passager. Comme elles sont autant positives que négatives, elles n'introduisent pas de biais systématiques. Finalement, étant donné ces faibles différences, il est préférable d'utiliser la loi normale pour des raisons de robustesse. Une distribution empirique tirée d'une autre situation présenterait probablement le même ordre de différence par rapport à la normale mais il est peu probable qu'elle présente les mêmes différences.

La seconde hypothèse porte sur la forte corrélation entre la demande journalière des paires O/D susceptibles d'utiliser les mêmes vols. Nous avons d'abord calculé la corrélation entre des groupes de paires O/D d'une même région.

L'Ouest Canadien, le triangle Montréal-Ottawa-Toronto, et les liaisons de Montréal ou Toronto vers les grandes villes Américaines ont d'abord été étudiés. Les coefficients de corrélations entre ces O/D ou les déplacements d'affaires sont importants, sont de l'ordre de .8 et .9 . Pour les origines Montréal et Toronto et les destinations dans l'Ouest Canadien où les distances sont plus grandes, les corrélations sont de l'ordre de .7 . Entre les marchés de l'Atlantique nord, les corrélations sont plutôt de l'ordre de .5 .

Dans ces conditions, l'utilisation d'un coefficient de corrélation $\rho$ égale à 1 pour l'estimation des $\sigma$, produit une surestimation moyenne de 1.5 . En réduisant $\rho$ à .7, on fait disparaître cette surestimation mais l'erreur sur l'assignation des passagers n'est réduite que de 11.3 à 11.2 . De même, en utilisant $\rho = .9$ pour les arcs courts, $\rho = .7$ pour les arcs moyens et $\rho = .5$ pour les arcs longs, l'erreur reste égale à 11.1 . Ces améliorations ne sont pas assez significatives pour qu'il soit justifié de les intégrer définitivement au modèle .

Donc, l'hypothèse $\rho = 1$ peut être utilisée en première approximation. Pour obtenir de meilleurs résultats, il faudrait développer des méthodes beaucoup plus compliquées pour estimer les $\sigma$.

## ANNEXE 2

### Etude de la convexité de la fonction économique

La convexité de la fonction économique est importante car elle nous assure l'optimalité globale des solutions obtenues avec les algorithmes. Sans la convexité, la solution trouvée peut être optimale mais rien ne le garantit.

Nous allons montrer que la fonction économique de ce modèle n'est pas toujours convexe et étudier des situations particulières la rendant convexe.

Le terme $Z_2$ étant linéaire, il s'ensuit que la convexité de la fonction objective dépend de celle du terme $Z_1$. Ce terme a été défini par l'équation (2.13)

$$Z_1 = \sum_r \sum_{q \in Q_r} B^r h_q \sum_{a \in F_q} \frac{E_a(h)}{f_a}$$

Analysons d'abord la composante $\frac{E_a(h)}{f_a}$. Rappelons d'abord qu'en (2.5) on a défini

$$E_a(h) = \sigma_a \left( f(p) - P[1 - F(p)] \right)$$

où p dépend de h et s'exprime comme

$$p = \frac{c_a - f_a}{\sigma_a}$$

D'autre part, se référant à (2.8) on peut exprimer $\sigma_a$ comme

$$\sigma_a = \sum_{q \in E_a} \frac{h_q}{u_{r_q}} \bar{\sigma}_{r_q}$$

où $u_{r_q}$ est le nombre total de passagers pour le marché $r_q$, et $\bar{\sigma}_{r_q}$ est l'écart-type pour le marché $r_q$ qu'on prend égal à l'écart-type pour la paire O-D à laquelle se rattache le marché $r_q$. Ainsi,

$$E_a(h) = \sum_{q \in E_a} \frac{h_q}{u_{r_q}} \bar{\sigma}_{r_q} \left(f(p) - p[1 - F(p)]\right) .$$

Voici une série de trois lemmes établissant la convexité de $E(p, 0, 1) = f(p) - p[1 - F(p)]$ .

Lemme 1: $\frac{d}{dp} f(p) = - p f(p)$

Preuve: $\frac{d}{dp} \left( \frac{1}{\sqrt{2\pi}} e^{-\frac{p^2}{2}} \right) = - p \frac{1}{\sqrt{2\pi}} e^{-\frac{p^2}{2}}$

Lemme 2: $\frac{d}{dp} E(p, 0, 1) = [F(p) - 1]$

Preuve: $\frac{d}{dp} \left(f(p) - p[1 - F(p)]\right)$

$= f'(p) - [1 - F(p)] + p f(p)$

$= [F(p) - 1]$ par lemme 1

Lemme 3: $E(p, 0, 1) = f(p) - p[1 - F(p)]$ est une fonction convexe de p

Preuve: $\frac{d^2}{dp^2} E(p, 0, 1) = \frac{d}{dp} \frac{d\, E(p, 0, 1)}{dp} = \frac{d}{dp} [F(p) - 1]$ (lemme 2)

$= f(p) > 0$

Même si $E(p, 0, 1)$ est une fonction convexe la fonction $E_a(h) = \sigma_a \cdot E(p(h), 0, 1)$ n'est pas toujours convexe. Voici un contre exemple l'illustrant:

Soient trois routes avec $h_1$, $h_2$, $h_3$ leurs nombres de passagers. Supposons que l'arc a est utilisé par les routes 1 et 2, i.e. $E_a = \{1, 2\}$. Soient $\lambda_1 = \frac{\bar{\sigma}_{r_1}}{u_{r_1}} = .4$ et $\lambda_2 = \frac{\bar{\sigma}_{r_2}}{u_{r_2}} = .1$ . La capacité de l'avion desservant l'arc a est de 100 places. Le tableau 9 évalue $E_a(h)$ aux trois points $(h_1^1, h_2^1, h_3^1) = (45, 50, 5)$, $(h_1^2, h_2^2, h_3^2) = (90, 0, 10)$ et $\frac{1}{2}(h_1^1, h_2^1, h_3^1) + \frac{1}{2}(h_1^2, h_2^2, h_3^2) = (67.5, 25, 7.5)$

TABLEAU 9

POINTS ILLUSTRANT LA NON-CONVEXITE DE $E_a(h)$

| h | (45, 50, 5) | (67.5, 25, 7.5) | (90, 0, 10) |
|---|---|---|---|
| a | 23 | 29.5 | 36 |
| p | .22 | .25 | .28 |
| $E(p, 0, 1)$ | .299 | .287 | .274 |
| $E_a(h)$ | 6.88 | 8.46 | 9.75 |

Ces points illustrent bien que $E_a(h)$ est non-convexe car $\frac{1}{2}(6.88) + \frac{1}{2}(9.75) = 8.32 < 8.46$ . Par contre, la courbure négative n'est pas très prononcée. De plus, si on change les paramètres $\lambda$,c ou si on change les trois points, la courbure deviendra probablement positive. Il est à noter que beaucoup d'effort a été déployé pour obtenir ce contre exemple et que c'est le plus concave que nous ayons pu trouver.

Dans le cas particulier où les coefficients de variation sont identiques pour toutes les O-D (i.e. $\frac{\sigma_r}{u_r} = \lambda$ pour tout r) nous allons montrer que $E_a(h)$ et $\frac{E_a(h)}{f_a}$ sont des fonctions convexes. Dans ce cas, on peut exprimer $E_a(h)$ comme une fonction de $f_a$ car,

$\sigma_a = \sum_{q \in E_a} h_q \lambda = f_a \lambda$ et $p = \frac{c_a - \sum_{q \in E_a} h_q}{\sigma_a} = \frac{c_a - f_a}{f_a \lambda}$. Comme

$(\forall q \in E_a)\ \frac{d}{d h_q} f_a = 1$, il suffit de montrer que $E_a(h)$ est une fonction convexe de $f_a$ pour établir la convexité par rapport à h. Pour simplifier la notation, posons $E_a(p, 0, 1) = G(p)$; ainsi, nous avons $E_a(f_a) = \lambda f_a\ G(p(f_a))$.

Lemme 4: $E_a(f_a)$ est une fonction convexe de $f_a$ dans le cas où tous les coefficients de variation sont égaux.

Preuve: $\frac{d\,p}{d\,f_a}(f_a) = \frac{d}{d\,f_a}\left(\frac{c_a - f_a}{\lambda f_a}\right) = \frac{-\lambda f_a - (c_a - f_a)\lambda}{(\lambda f_a)^2}$

$= \frac{-\lambda f_a - c_a \lambda + \lambda f_a}{(\lambda f_a)^2} = \frac{-c_a}{\lambda f_a^2}$.

$\frac{d\,E(f_a)}{d\,f_a} = \frac{d}{df_a}\left(\lambda f_a\ G(p(f_a))\right)$

$= \lambda\, G(p(f_a)) + \lambda f_a \frac{d}{dp} G(p(f_a)) \frac{d}{df_a}\left(\frac{c_a - f_a}{\lambda f_a}\right)$

$$= \lambda\, G(p(f_a)) + \lambda f_a [F(p(f_a)) - 1] \left( \frac{-c_a}{\lambda f_a^2} \right)$$

$$= \lambda\, G(p(f_a)) - \frac{c_a}{f_a} [F(p(f_a)) - 1] .$$

$$\frac{d^2 E(f_a)}{df_a^2} = \frac{d}{df_a} \left\{ \lambda\, G(p(f_a)) - \frac{c_a}{f_a} [F(p(f_a)) - 1] \right\}$$

$$= \lambda \frac{d}{dp} G(p(f_a)) \left( \frac{-c_a}{\lambda f_a^2} \right) - \frac{c_a \frac{d}{dp} [F(p(f_a)) - 1](-c_a/\lambda f_a^2) f_a - c_a [F(p(f_a)) - 1]}{f_a^2}$$

$$= \lambda [F(p) - 1] \left( \frac{-c_a}{\lambda f_a^2} \right) + \frac{c_a^2}{\lambda f_a^3} f(p(f_a)) + \frac{c_a}{f_a^2} [F(p(f_a)) - 1]$$

$$= \frac{c_a^2}{\lambda f_a^3} f(p(f_a)) > 0 .$$

Lemme 5: $\frac{E_a(f_a)}{f_a}$ est une fonction convexe de $f_a$ si $f_a < c_a$ et si tous les coefficients de variation sont égaux et inférieurs à .4

Preuve: $\frac{E_a(f_a)}{f_a} = \frac{\lambda f_a\, G(p(f_a))}{f_a} = \lambda\, G(p(f_a))$

$$\frac{d}{df_a} \lambda\, G(p(f_a)) = \lambda \frac{d\, G}{dp} (p(f_a)) \cdot \frac{d\, p(f_a)}{d\, f_a} =$$

$$\lambda [F(p(f_a)) - 1] \left( \frac{-c_a}{\lambda f_a^2} \right)$$

$$\frac{d^2}{df_a^2}\frac{E(f_a)}{f_a} = \frac{\lambda\, d}{df_a}\left( [F(p(f_a)) - 1] \left(\frac{-c_a}{\lambda f_a^2}\right)\right)$$

$$= \lambda\, f(p(f_a)) \left(\frac{-c_a}{\lambda f_a^2}\right)^2 + \lambda[F(p(f_a)) - 1] \left(\frac{2c_a}{\lambda f_a^3}\right)$$

$$= \frac{c_a}{f_a^3}\left(\frac{c_a}{\lambda f_a} f(p(f_a)) - 2\,[F(p(f_a)) - 1]\right).$$

Si $p > 0$, alors on peut vérifier à l'aide d'une table que $[F(p) - 1] < 1.25\, f(p)$

Donc $\frac{d^2}{df_a^2}\frac{E(f_a)}{f_a} > \frac{c_a}{f_a^3} f(p(f_a)) \left(\frac{c_a}{\lambda f_a} - 2.5\right) \geq 0$ si $\frac{c_a}{\lambda f_a} \geq 2.5$.

Ainsi deux conditions suffisantes pour assurer la convexité de $\frac{E_a(f_a)}{f_a}$ sont $p > 0$ et $\frac{c_a}{\lambda f_a} \geq 2.5$. Or, $p > 0$ si et seulement si $f_a < c_a$. D'autre part, la condition $\frac{c_a}{\lambda f_a} \geq 2.5$ peut être remplacée par une condition plus stricte de la forme $\frac{f_a}{\lambda f_a} \geq 2.5$, i.e. $\lambda \leq 0.4$. Ceci complète la preuve du lemme.

En pratique, on a observé que les coefficients de variations se situent généralement entre .2 et .3 (donc ne sont pas très différents) et que $f_a < .9c_a < c_a$. Ainsi, on est loin de l'exemple pathologique et beaucoup plus près des conditions des lemmes 4 et 5 par conséquent, $E_a(f_a)$ et $E_a(f_a) / f_a$ sont probablement convexes.

Etudions maintenant la convexité de $Z_1$ quand les $\frac{E_a(h)}{f_a}$ sont convexes. Nous allons voir que même si $\frac{E_a(h)}{f_a}$ est une fonction convexe, le terme $B^r h_q \sum_{a\in F_q} \frac{E_a(h)}{f_a}$ tout en étant une fonction convexe de chacune des variables $h_q$ n'est pas une fonction convexe de la variable vectorielle $h = (h_q \mid q \in Q)$. Le lemme 6 établit la convexité de chacun des termes de $Z_1$ comme fonction d'une variable $h_q$.

Lemme 6: Soit q fixe, si pour tout $a \in F_q$, $\frac{E_a(h)}{f_a}$ est une fonction convexe de h alors, $B^r h_q \sum_{a\in F_q} \frac{E_a(h)}{f_a}$ est une fonction convexe de $h_q$

Preuve: Il est immédiat que les $\frac{E_a(h)}{f_a}$ et $\sum_{a\in F_q} \frac{E_a(h)}{f_a}$ sont des fonc- convexes de $h_q$ . D'autre part

$$\frac{d^2}{dh_q^2} B^r h_q \sum_{a\in F_q} \frac{E_a(h)}{f_a}$$

$$= B^r \frac{d}{dh_q} \left( \sum_{a\in F_q} \frac{E_a(h)}{f_a} + h_q \frac{d}{dh_q} \sum_{a\in F_q} \frac{E_a(h)}{f_a} \right)$$

$$= B^r \left\{ \frac{d}{dh_q} \left( \sum_{a\in F_q} \frac{E_a(h)}{f_a} \right) + \frac{d}{dh_q} \sum_{a\in F_q} \frac{E_a(h)}{f_a} + h_q \frac{d^2}{dh_q^2} \sum_{a\in F_q} \frac{E_a(h)}{f_a} \right\} \geq 0$$

car $B^r > 0$, $\frac{E_a(h)}{f_a}$ croissante, $h_q \geq 0$ et $\frac{E_a(h)}{f_a}$ convexe.

Ceci complète la preuve.

Illustrons maintenant à l'aide d'un contre exemple que $B^r h_q \sum_{a \in F_q} \frac{E_a(h)}{f_a}$ et $Z_1$ ne sont pas des fonctions convexes de h. Considérons le réseau avec deux origines et une destination, représentés par la figure suivante.

FIGURE 38

RESEAU POUR LEQUEL LA FONCTION ECONOMIQUE N'EST PAS CONVEXE.

Les arcs pointillés représentent l'accès aux routes, $h_i$ est le nombre de passagers qui utilisent la route i. Les arcs $a_j$, $j = 1, \ldots, 4$ représentent les segments de vols. Le nombre de passagers utilisant l'arc $a_j$ est désigné, suivant la notation déjà définie, par $f_{a_j}$. Supposons que les coûts des billets sont de \$100 entre $O_1$ et D et de \$1 entre $O_2$ et D. Considérons une demande de 10 passagers entre $O_1$ à D et 30 entre $O_2$ à D. Pour chaque segment de vol, prenons un taux de passagers refusés, égal à $f^2_{a_j} / 1000$. Cette fonction est

convexe. Le terme $Z_1$ de la fonction économique est donné par

$$Z_1 = \frac{100\, h_1^3}{1000} + 100\, h_2 \frac{\left(h_2^2 + (h_2 + h_3)^2\right)}{1000} + h_3 \frac{(h_2 + h_3)^2}{1000} + \frac{h_4^3}{1000}$$

Le tableau suivant montre que $Z_1$ et en particulier son deuxième terme n'est pas convexe lorsqu'on considère les trois points suivants $(h_1^1, h_2^1, h_3^1, h_4^1) = (10, 0, 30, 0)$, $(h_1^2, h_2^2, h_3^2, h_4^2) = (0, 10, 0, 30)$ et $\frac{1}{2}(h_1^1, h_2^1, h_3^1, h_4^1) + \frac{1}{2}(h_1^2, h_2^2, h_3^2, h_4^2) = (5, 5, 15, 15)$

TABLEAU 10

POINTS ILLUSTRANT LA NON-CONVEXITE DE $Z_1$

| | (10,0,30,0) | (0,10,0,30) | (5,5,15,15) |
|---|---|---|---|
| $Z_1$ | 127 | 227 | 234 |
| $.1 \times h_2(h_2^2 + (h_2+h_3)^2)$ | 0 | 200 | 212.5 |

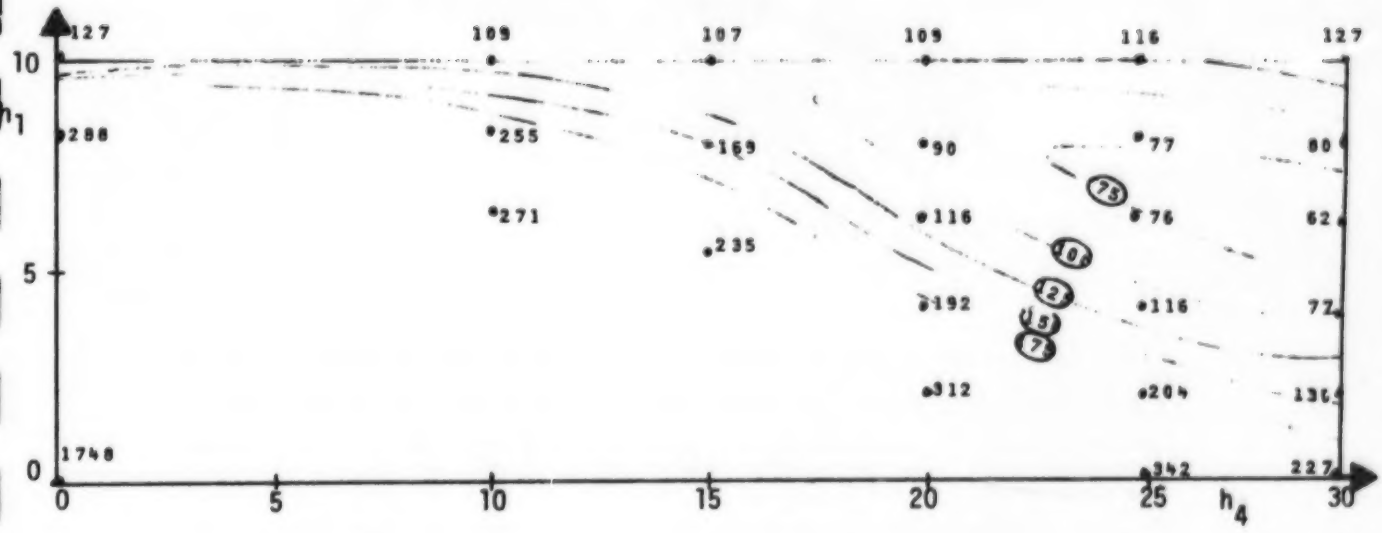
La figure 39 donne une vue d'ensemble du comportement de cette fonction non-convexe. Son domaine de dimension 4 a été projeté dans $R^2$ de façon à satisfaire les équations de la demande:

$h_1 + h_2 = 10$, $h_3 + h_4 = 30$.

Ainsi $h_2 = 10 - h_1$ et $h_3 = 30 - h_4$ .

FIGURE 39

COURBE DE NIVEAU D'UN EXEMPLE
DE FONCTION ECONOMIQUE NON-CONVEXE




Cet exemple permet aussi de montrer que la fonction $Z_1$ n'est même pas pseudo-convexe ni quasi-convexe. Il est à remarquer que même si cette fonction n'est pas convexe, elle ne comporte pas d'optimum local qui n'est pas un optimum global.

Nous allons maintenant étudier un cas particulier où $Z_1$ est convexe.

Lemme 7: Si les coûts des passagers refusés faute de place ne dépendent pas de leur paire O-D mais, seulement du segment de vol où ils sont refusés et si les $E_a(h)$ sont convexes alors, $Z_1$ est une fonction convexe.

Preuve: $$Z_1 = \sum_r \sum_{q \in Q_r} h_q \sum_{a \in F_q} B^a \frac{E_a(h)}{f_a}$$

$$= \sum_a B^a \left( \sum_{q \in E_a} h_q \right) \frac{E_a(h)}{f_a}$$

$$= \sum_a B^a f_a \frac{E_a(h)}{f_a}$$

$$= \sum_a B^a E_a(h)$$

Ainsi par le lemme 4 et le lemme 7 si tous les coefficients de variations sont égaux et si les coûts des passagers refusés ne dépendent que de l'arc, la fonction $Z_1$ est convexe. En pratique, il est probable que $Z_1$ soit convexe car les coûts des billets des passagers utilisant un même avion ne sont pas très différents, sûrement pas aussi différents que dans l'exemple pathologique illustrant une fonction $Z_1$ non-convexe.

Au cours de l'expérimentation avec le réseau d'Air Canada, une cinquantaine de résolutions en variant les paramètres du problème ont donné des solutions stables tant du point de vue des variables que de l'objectif. Ceci semble indiquer que le problème d'optimum local n'a pas été rencontré.

## ANNEXE 3

### Nombre de passagers excédentaires exprimé en fonction du nombre de passagers trouvant place dans l'avion

Lors de la définition du modèle P3, nous avons utilisé $P_q$ comme fonction de k les nombres de passagers trouvant place dans les avions. Comme nous l'avons mentionné à ce moment-là, $P_q$ a été défini à l'équation (2.11) en fonction des nombres de passagers tentant d'utiliser chacune des routes. Voyons maintenant comment exprimer $P_q$ comme fonction de k à partir de sa définition en fonction de h.

Rappelons la définition de $P_q$ donnée en (2.11)

$$P_q = h_q \sum_{a \in F_q} \frac{E_a(h)}{f_a} \quad \text{où} \quad f_a = \sum_{q \in E_a} h_q \ .$$

Rappelons aussi que $h_q$ le nombre de passagers tentant d'utiliser la route q est relié à $k_q$ le nombre de passagers utilisant la route q par (2.17) $h_q = k_q + P_q$ . De plus $f_a = k_a + E_a(h)$ en posant $k_a = \sum_{q \in E_a} k_q$ . En utilisant ces relations, (2.11) devient

$$P_q = (k_q + P_q) \sum_{a \in F_q} \frac{E_a(h)}{k_a + E_a(h)}$$

et en isolant $P_q$

$$P_q = k_q \frac{\sum_{a \in F_q} \left(E_a(h)/(k_a + E_a(h)\right)}{1 - \sum_{a \in F_q} \left(E_a(h)/(k_a + E_a(h))\right)} \ .$$

Ainsi, le problème se réduit à exprimer $E_a$ en fonction de k .

Il n'est pas possible d'obtenir l'expression de $E_a$ comme fonction de k de la même façon en substituant (2.17) dans l'équation (2.9) définissant $E_a$ . En effet, il faudrait résoudre une équation de degré 7 pour isoler $E_a(k)$ .

A défaut de l'expression exacte une approximation peut être obtenue de la façon suivante. D'abord l'équation (2.9) a été simplifiée. La Figure 40 présente différentes variantes considérées. Premièrement $y_0$ est la fonction exacte définie par la relation (2.5)

$$y_0(f_a) = \lambda_a f_a \left(f(p) - p(1 - F(p))\right) \text{ avec } p = \frac{c_a - f_a}{\lambda_a f_a} .$$

Deuxièmement $y_1$ est la fonction définie par l'équation (2.9)

$$y_1(f_a) = \lambda_a f_a \left(\frac{5 - p}{5}\right)^7 \times .4 \quad \text{avec} \quad p = \frac{c_a - f_a}{\lambda_a f_a} .$$

Troisièmement $y_2$ est une simplification où $\sigma = \lambda_a f_a$ est remplacé par une constante soit $\lambda_a \cdot 8\, c_a$

$$y_2(f_a) = .32\, \lambda_a c_a \left(1 - \frac{(c_a - f_a)}{4 \lambda_a c_a}\right)^7 .$$

Quatrièmement $y_3$ s'obtient de $y_2$ en augmentant l'exposant pour compenser partiellement la réduction de courbure qu'entraîne l'utilisation du $\sigma$ constant

$$y_3(f_a) = .32\, \lambda_a c_a \left(1 - \frac{(c_a - f_a)}{4 \lambda_a c_a}\right)^{8.67} .$$

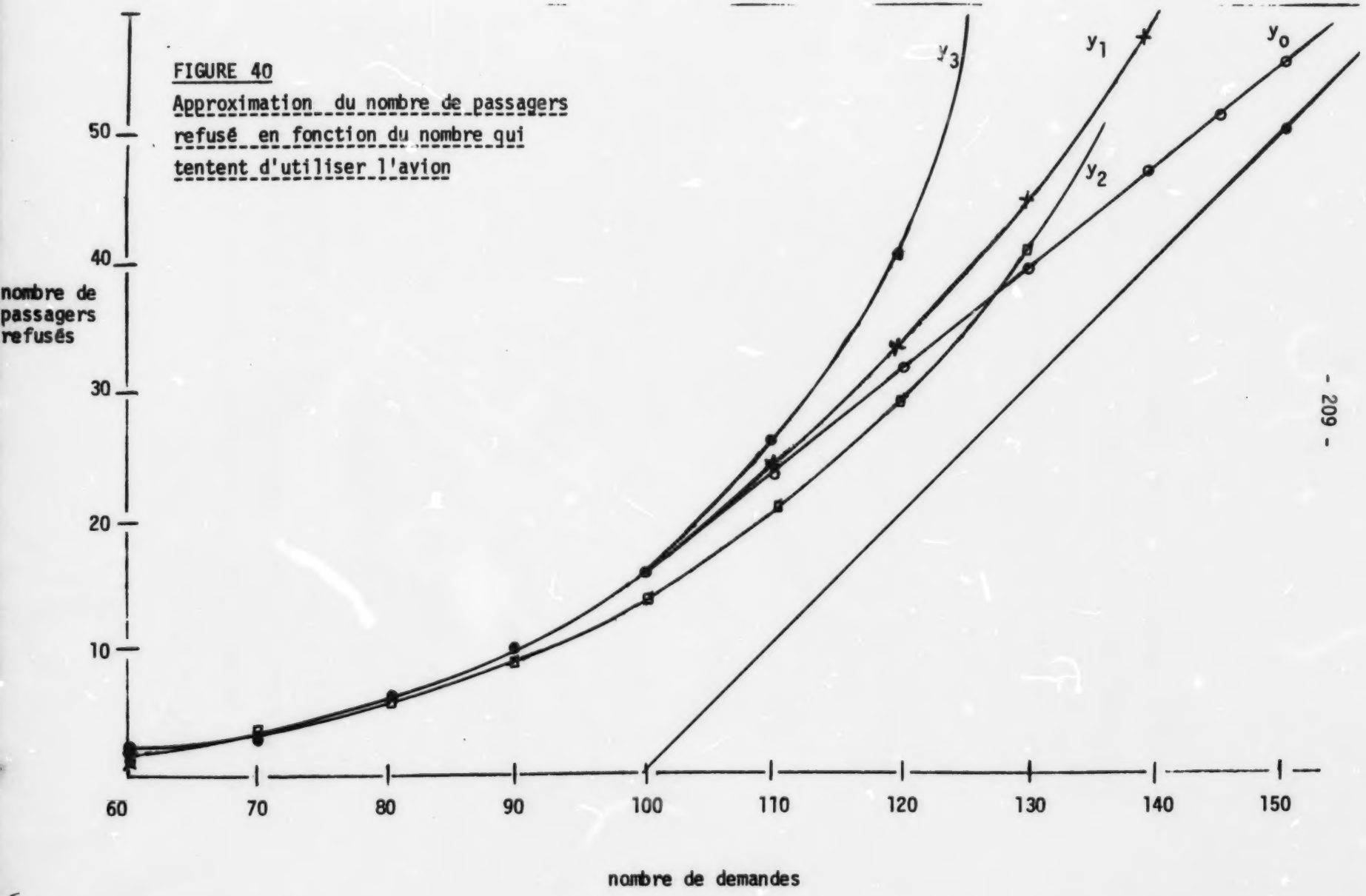



FIGURE 40
Approximation du nombre de passagers refusé en fonction du nombre qui tentent d'utiliser l'avion

L'examen de la Figure montre que $y_3$ est une très bonne approximation de $y_o$ pour $0 \leq f_a \leq c_a$ . L'erreur dans la région $f_a > c_a$ n'a pas d'importance, car cette section de la courbe ne sera pas utilisée. Exprimons maintenant $y_3$, comme fonction de $k_a$ dans le cas où $f_a \leq c_a$ . A cause de la complexité de $y_3$ il n'était pas possible de substituer (2.17) et d'isoler $y_3(k_a)$ . Développons plutôt $y_3(f_a)$ en série de Taylor et conservons seulement les trois premiers termes pour y substituer (2.17) et isoler $y_3(k_a)$ . La formule obtenue est

$$y_3(k_a) = \frac{.32\ \lambda_a\ c_a\ \left((1 - 4\ \lambda_a)\ c_a + k_a\right)^{8.67}}{(4\lambda_a\ c_a)^{8.67} - 2.77\ \lambda_a\ c_a\ \left((1 - 4\ \lambda_a)\ c_a + k_a\right)^{7.67}}$$

pour $k_a \leq (1 - .4\lambda_a)\ c_a$ .

Pour $k_a > (1 - .4\ \lambda_a)\ c_a$ qui correspond à $f_a > c_a$ nous utilisons une approximation linéaire. Cette droite coïncide avec la fonction théorique en $f_a = c_a$ et $f_a = c_a + \sigma_a$ .

Son équation est la suivante

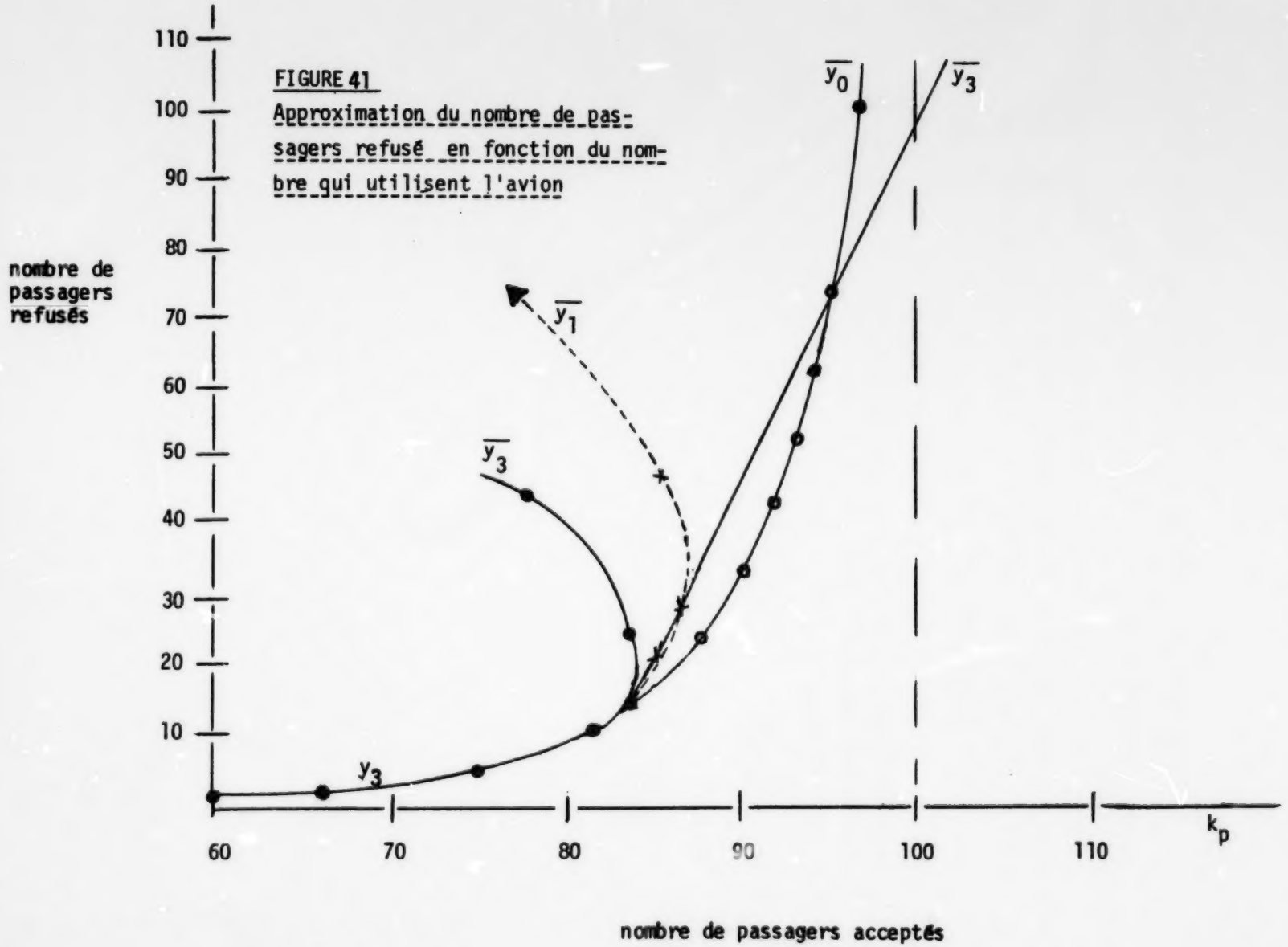
$$y_3(k_a) = \frac{.67 + .4\ \lambda_a}{.33 - .4\ \lambda_a}\ \left(h_a - (1 - .4\ \lambda_a)\ c_a\right)$$

pour $k_a > (1 - .4\ \lambda_a)\ c_a$ .

La Figure 41 présente l'approximation de $E_a(k_a)$ donnée par $y_3(k_a)$ . $\bar{y}_o$ représente les valeurs théoriques qui ont été calculées point par point. De même $\bar{y}_1$ et $\bar{y}_3$ correspondent à l'inversion point par point de $y_1(f_a)$ et $y_3(f_a)$ pour $f_a > c_a$ .

FIGURE 41

Approximation du nombre de passagers refusé en fonction du nombre qui utilisent l'avion



Le fait que $E_a(k_a)$ ne présente pas d'asymptote verticale quand $k_a$ s'approche de $c_a$ tout comme la courbe théorique est un avantage en pratique. Si l'hypothèse qu'il y a suffisamment de place pour que tous les passagers puissent voyager n'était pas réalisée, le modèle théorique serait insoluble. Avec $E_a(k)$ ainsi définie le problème reste soluble même si $k_a > c_a$ . Une telle solution indiquera que pour ce marché il n'y a pas suffisamment de place. Le reste de la solution sera valable et ainsi nous aurons toute l'information pertinente pour modifier cet horaire insatisfaisant.

## BIBLIOGRAPHIE

AGARD, J., SUDAROVICH, J. et HEMMER, F. (73)
Sélection et affectation optimale d'une flotte d'avion, R.A.I.R.O., Vol. 2, pp (3 - 26).

AGIN, N. (72)
Optimal Aircraft Scheduling, Proceeding NATO Conference, Sandefjord, Norway.

BOURQUE, M., FERLAND, J.A., et ROUSSEAU, J.-M. (76)
Itineraries Generator for Linear Network, Centre de recherche sur les transports, Université de Montréal, Rapport No 36.

CAMBRIDGE SYSTEMATICS INC (75)
Demand Model Development, Contract No Dot-OS-60061, Department of Transportation.

CHAN, Y.P. (72)
Route Network Improvement in Air Transportation Schedule Planning, Ph.D. Thesis, M.I.T., Flight Transportation Laboratory, Cambridge, Mass.

CHARNES, A. et COOPER, W. (61)
Multicopy Traffic Network Models, Theory of Traffic Flow(R. Herman editor), Elsevier, Amsterdam (1961).

COHEEN, M.S. (77)
A Decomposition Principle for Single Constraint Quadratic Programs, Technical Report Sol 77 - 2, System Optimization Laboratory, Stanford University.

DAFERMOS, C. (71)
An Extended Traffic Assignment Model with Application to Two-Way Traffic, Transportation Science, Vol. 5, No 4, pp (366 - 389).

DAIGLE, A. (76)
Modification et implantation d'un modèle linéaire pour une planification à moyen terme d'une flotte d'avions, Centre de recherche sur les transports, Université de Montréal, Rapport No 48.

ELCE, I. (70)
The Development and Implementation of Air Canada's Long Range Planning Model, AGIFORS Proceedings 1970.

ERIKSEN, S.E. et TENEJA, N.K. (75)
Unified Performance Measure of Airline Service, TIMS/ORSA Meeting, Las Vegas.

GAGNON, G. (67)
A Model for Flowing Passengers over Airline Network, Transportation Science, Vol. 1, No 3, pp (232 - 248).

GUNN, A. (64)
Airline System Simulation, Operation Research, Vol. 12, No 2, pp (206 - 229).

KARAMARDIAN, S. (69)
The Nonlinear Complementarity Problem with Application, JOTA, Vol. 4, No 2, pp (87 - 98).

KENNINGTON, J. (77)
A Survey of Linear Cost, Multicommodity Network Flows, TIMS/ORSA Meeting, Atlanta.

KINGSLEY, D. (68)
Airline Operation and Planning Model, Mc Donald Douglas, Long Beach, California, Paper 5058.

LUENBERGER, D.C. (73)
Introduction to Linear and Non Linear Programming, Addison Wesley, Reading, Mass.

MINOUX, M. (75)
Résolution des problèmes de multiflots en nombres entiers dans les grands réseaux, R.A.I.R.O., Vol. 3, pp (21 - 40).

MURCHLAND, J.D. (69)
Road Network Traffic Distribution in Equilibrium, Conference: Mathematical Methods in the Economics Sciences, Mathematisches Forschungsinstitut, Oberwolfach.

NGUYEN, S. (74)
Traffic Assignment Problem, An Algorithm for the Transportation Science, Vol 8, No 3, pp (203 - 216).

ORTEGA, J.M., RHEINBOLDT, W.C. (70)
Iterative Solution of Non Linear Equations in Several Variables, Academic Press.

POLLACK, M. (73)
An Inter-Active Computer Model for Airline Frequency Planning, CTR 1660, Commercial Transportation Research Department, Lockeed-California Company.

PRICE, CURRIE, DICKEY et DUCLOS (74)
Forcasting Traffic in an Air Transportation Network, TIMS/ORSA Meeting, Porto Rico.

RICHARDSON, J. (76)
An Optimization Approach to Routing Aircraft, Transportation Science, Vol. 10, No 1, pp (52 - 71).

ROCKAFELLAR, R.T. (70)
Convex analysis, Princeton University Press, Princeton, New Jersey.

RUBIN, J. (71)
A Technique for the Solution of Massive Set Covering Problems, with Application to Airline Crew Sceduling, IBM Technical Report # 320, 3004.

SCARBROUGH, J. (77)
A Mixed Integer Programming Model for the Scheduling of Aircraft by Time of Day within a Network of Airports, TIMS/ORSA Meeting, San Francisco.

SIMPSON, R. (69)
Scheduling and Routing Models for Airline Systems, M.I.T. Flight Transportation Lab., Cambridge, Mass.

SWANN, W. (72)
The Complete FA-4 Memo, M.I.T. Flight Transportation Lab., Cambridge, Mass.

VARGA, M. (62)
Matrix Iterative Analysis, Prentice Hall.

WARDROP, J.G. (52)
Some Theorical Aspects of Road Traffic Research, Proc. Inst. Civil Engeneers, Part II, Vol. 1, pp (325 - 378).

ZANGWILL, W. (69)
Nonlinear Programming: A Unified Approach, Prentice Hall Englewood Cliffs, N.J.